REMARKS

Claims 6–9, 26, 27 and 32-34 are pending in this application. Claim 34 is amended herein to clarify and more particularly define the invention. No new matter is added by this amendment. In light of the following amendments and remarks, applicants respectfully request reconsideration of this application and allowance of the claims to issue.

I. Rejection under 35 U.S.C. § 112

A. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 112, first paragraph, for allegedly failing to comply with the written description requirement.

Claim 6 as presented herein encompasses a specific genus of nucleic acid sequences encoding peptides immunochemically reactive with antibodies to the Epstein Barr Virus (EBV) VCA-p18 or VCA-p40 proteins, comprising an epitope of the VCA-p18 or VCA-p40 protein, encoded within the EBV open reading frames BFRF3 and BdRF1, respectively and wherein said antibodies are antibodies having the same reactivity with VCA-p18 as antibodies produced by the hybridomas deposited at the European Collection of Animal Cell Cultures under deposit numbers 93020413 or 93020412 or antibodies having the same reactivity with VCA-p40 as antibodies produced by the hybridoma deposited at the European Collection of Animal Cell Cultures under deposit number 93020414.

In claim 6, the nucleic acid sequences encode peptides comprising an epitope of the VCA-p18 or VCA-p40 protein. The peptides encoded by the nucleic acid sequences of claim 6 are further defined by their immunoreactivity with hybridoma-derived antibodies to VCA-p18 and VCA-p40 defined as European Collection of Animal Cell Cultures deposit numbers 93020413 or 93020412 (VCA-p18) or deposit number 93020414 (VCA-p40). As one of skill in the art would recognize, hybridoma derived antibodies are monoclonal. One of skill in the art further recognizes that monoclonal antibodies are specific to a single epitope and that all of the antibodies produced from a single hybridoma are identical. Therefore, contrary to the arguments in the Office Action,

these antibodies are not heterogeneous and, thus, are sufficient to define the genus of nucleic acid sequences encompassed by claim 6.

Furthermore, the specification demonstrates actual reduction to practice of the nucleic acids of claim 6 and in particular, provides several examples of peptides comprising epitopes of this invention (e.g., SEQ ID NOs:2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22; see pages 7-13 for description of peptides and fragments of this invention and Examples 4 and 5, Figures 4-6 and Table 1) that are reactive with the EBV VCA-p18 or VCA-p40 monoclonal antibodies of claim 6 and also provides examples of nucleic acid sequences encoding such peptides (e.g., SEQ ID NO:1, SEQ ID NO:3). Accordingly, one of skill in the art would recognize that applicants were in possession of the nucleic acid sequences of claim 6 at the time the present application was filed, as evidenced by the large numbers of representative species disclosed in the specification.

Therefore, applicants respectfully submit that all of the members of the genus of nucleic acids of claim 6 are adequately defined both structurally and functionally, leading one of ordinary skill in the art to the reasonable conclusion that applicants were in possession of the invention of claim 6 at the time this application was filed.

Claims 7 and 8 as presented herein respectively encompass a specific genus of nucleic acid sequences comprising the nucleotide sequence or a subsequence of SEQ ID NO:1, wherein the subsequence encodes a peptide that comprises an epitope that is immunochemically reactive with antibodies to EBV VCA-p19 protein (claim 7), and a specific genus of nucleic acid sequences comprising the nucleotide sequence or a subsequence of SEQ ID NO:3, wherein the subsequence encodes a peptide that comprises an epitope that is immunochemically reactive with antibodies to EBV VCA-p40 protein (claim 8).

In both claims 7 and 8, the subsequence is defined as encoding a peptide comprising an epitope. It would be readily recognized by one of skill in the art that applicants were in

possession of the genus of nucleic acid sequences of claims 7 and 8 at the time the application was filed because the specification provides several examples of peptides comprising epitopes of this invention (e.g., SEQ ID NOs: 2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22; see pages 7-13 for description of peptides and fragments of this invention and Examples 4 and 5, Figures 4-6 and Table 1) that are reactive with the EBV VCA-p18 or VCA-p40 antibodies of this invention and also provides examples of nucleic acid sequences encoding such peptides (e.g., SEQ ID NO:1, SEQ ID NO:3). Thus, one of skill in the art would recognize that applicants were in possession of the nucleotide sequences and subsequences of claims 7 and 8 at the time the present application was filed, as evidenced by the large numbers of representative species disclosed in the specification.

In particular, all of the members of the genus of the nucleotide sequences of claims 7 and 8 could be readily identified by one of ordinary skill in the art on the basis of the disclosure of the nucleotide sequences of SEQ ID NO:1 or SEQ ID NO:3. Such a genus is not overly broad, considering that every member must be a subsequence of a disclosed sequence (SEQ ID NO:1 or SEQ ID NO:3), thereby defining the members of the genus structurally AND every member of the genus must also meet the functional requirement of encoding an EBV peptide comprising an epitope that is immunochemically reactive with antibodies to the EBV VCA-p18 protein or the EBV VCA-p40 protein. Thus, all of the members of the genus of nucleic acids of claims 7 and 8 are adequately defined both structurally and functionally, leading one of ordinary skill in the art to the reasonable conclusion that applicants were in possession of the invention of claims 7 and 8 at the time this application was filed.

Claims 9, 26 and 27 depend from claims 6, 7 and 8, respectively, and recite a vector molecule comprising the nucleic acid molecule of each respective independent claim. Because the nucleic acid sequences of claims 6, 7 and 8 are adequately described in the specification, the vectors of these claims are adequately described as well.

With regard to claims 32-33, the specification presents data that demonstrate that the inventors produced more than 330 12 mers of VCA-p40 and more than 160 12 mers of VCA-p18 as described in Examples 4 and 5 and as shown in Figures 4-6 and in Table 1 of the specification, thereby adequately describing the genus of 12 contiguous amino acids as set forth in these claims.. Specifically, Example 4 describes the production of a full set of peptides with a length of 12 amino acids and an overlap of 11 amino acids of the amino acid sequences of both ORFs BFRF3 (VCA-p18) and BdRF1 (VCA-p40) (page 30). These peptides were assayed for immunoreactivity with EBV-specific antibodies (Example 4, page 31 and Example 5, page 33) and results of these assays are shown for the VCA-p18 peptides in Figures 4 and 5 and for the VCA-p40 peptides in Figure 6. Specifically, Figure 6 shows immunoreactivity results of almost 340 peptides of VCA-p40 and Figures 4 and 5 show such results for more than 160 peptides of VCA-p18. Thus, one of skill in the art would reasonably conclude that the peptides of claims 32 and 33 are adequately supported in the specification

Further, claim 34 as presented herein recites an isolated nucleic acid sequence encoding the amino acid sequence of SEQ ID NO:5 or SEQ ID NO:6 or a combination of both, wherein said amino acid sequence is immunochemically reactive with antibodies to the Epstein-Barr Virus VCA-p18 protein. Thus, claim 34 recites a nucleic acid sequence encoding the specific amino acid sequences of SEQ ID NO: 5 and/or SEQ ID NO:6, which are disclosed in the specification at least on page 9, second paragraph. Thus, the nucleic acid sequences of claim 34 are adequately supported in the specification.

Thus, at least for the reasons set forth above, applicants believe that this rejection has been overcome and its withdrawal and allowance of the pending claims are respectfully requested.

B. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 112, first paragraph, for allegedly failing to comply with the enablement requirement. Specifically, the Action states that to the extent that the claimed sequences are not adequately

described in the instant disclosure, claims 6-9, 26, 27 and 32-34 are also rejected under 35 U.S.C. § 112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to make and/or use the invention.

As discussed above, the subject matter of claims 6-9, 26, 27 and 32-34 is adequately described in the present specification. The specification not only adequately discloses the full genus of nucleic acid sequences of this invention, but also provides detailed teachings of how to make and use these nucleic acid sequences, See, in particular, the Examples set forth on pages 22-26, wherein numerous working examples are provided of the production and testing of numerous peptides of this invention. Thus, applicants respectfully submit that the present invention is adequately enabled and applicants thereby respectfully request withdrawal of this rejection.

II. Rejection under 35 U.S.C. § 102(b)

A. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Laux et al. (*EMBO J.* 7:769-774 (1988)). Specifically, the Action states that Laux et al. teaches a nucleic acid sequence comprising instant SEQ ID NO:1 which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ ID NO:5). The Action further states that Laux et al. teaches a nucleic acid sequence comprising a sequence that shares 98.8% homology with instant SEQ ID NO:3 (subsequence thereof), which encodes 12 contiguous amino acids of an EBV VCA-40. On this basis, the Action concludes that Laux et al. anticipates the instant claims. Applicants respectfully disagree and traverse this rejection.

Specifically, applicants have performed multiple alignments comparing both the nucleotide and amino acid sequences of the present invention with those of Figure 2 of Laux et al. (NCBI Accession No. Y00835.1) and the sequence homology asserted in the Action to be present was not found (See enclosed Alignments 1-6). Accordingly, applicants respectfully submit that Laux et al. fails to disclose the nucleotide sequences of SEQ ID NO:1, or SEQ ID

NO:3 or any subsequences thereof encoding 12 contiguous amino acids, as claimed herein. If the Examiner maintains this rejection, it is respectfully requested that the Examiner specifically point out what portion of the sequences of Laux et al. have homology with the sequences of the present invention. Otherwise, applicants respectfully request that this rejection be withdrawn.

B. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Bankier et al. (*Mol. Biol. Med.*. 1:425-445 (1983)). Specifically, the Action states that Bankier e al. teaches a nucleic acid sequence comprising instant SEQ ID NO:1, which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ ID NO:5). The Action further states that Bankier et al. teaches a nucleic acid sequence comprising a sequence that shares 98.8% homology with instant SEQ ID NO:3 (subsequence thereof), which encodes 12 contiguous amino acids of an EBV VCA-40. On this basis, the Action concludes that Bankier et al. anticipates the instant claims. Applicants respectfully disagree and traverse this rejection.

Specifically, applicants have performed multiple alignments comparing both the nucleotide and amino acid sequences of the present invention with those of Figure 2 of Bankier et al. and the sequence homology asserted in the Action to be present was not found (See enclosed Alignments 7-38). Accordingly, applicants respectfully submit that Bankier et al. fails to disclose the nucleotide sequences of SEQ ID NO:1, or SEQ ID NO:3 or any subsequences thereof encoding 12 contiguous amino acids as claimed herein. If the Examiner maintains this rejection, it is respectfully requested that the Examiner specifically point out what portion of the sequences of Bankier et al. have homology with the sequences of the present invention. Otherwise, applicants respectfully request that this rejection be withdrawn

The points and concerns raised in the outstanding Office Action having been addressed in full, it is respectfully submitted that all of the claims of this application are in condition for allowance, which action is respectfully requested. Should the Examiner have any remaining

concerns, the Examiner is invited and encouraged to contact the undersigned attorney directly by telephone in order to expedite the prosecution of this application to allowance.

The Commissioner is authorized to charge Deposit Account No. 50-0220 in the amount of $120.00 as fee for a one-month extension of time. This amount is believed to be correct. However, the Commissioner is authorized to charge any deficiency or credit any overpayment to Deposit Account No. 50-0220.

Respectfully submitted,

Mary L. Miller
Registration No. 39,303

Customer No. 20792
Myers Bigel Sibley & Sajovec, P.A.
P. O. Box 37428
Raleigh, North Carolina 27627
Telephone: (919) 854-1400
Facsimile: (919) 854-1401

CERTIFICATE OF EXPRESS MAILING UNDER 37 CFR 1.10
"Express Mail" mailing label number: EV887524629US
Date of Deposit: July 23, 2007
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Tracy Wallace

DNA Strider 1.4f6 ### Friday, July 13, 2007 4:44:32 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 1.xdna x Laux et al. EBV terminal gene.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

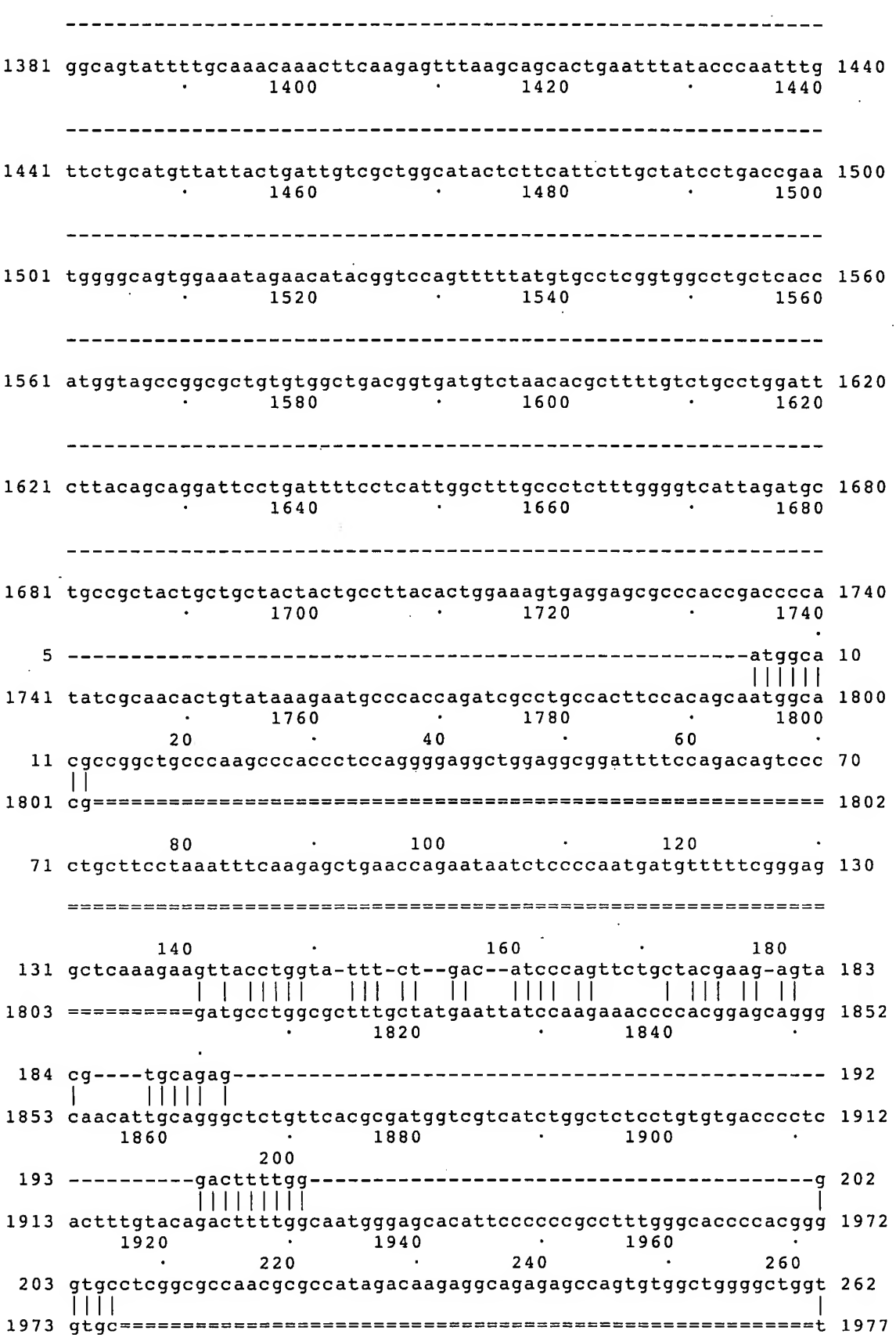
DNA sequence 2227 bp gcagtgtgtgaa ... aaaaaaaaaaaa linear

Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 1. Comparison of nucleotide sequence of SEQ ID NO:1 with the nucleotide sequence of Fig. 2 of Laux et al.

```
   ------------------------------------------------------------

     1 gcagtgtgtgaagattgtcacagctgctggtttggagaaaacgggggtgggcggtgatca 60
               .         20          .         40          .         60

   ------------------------------------------------------------

    61 gggagaacaattccccggggacacctgcacgagacccctgggctctcaggaactccgccc 120
               .         80          .        100          .        120

   ------------------------------------------------------------

   121 aggtcttgccaattggggtgatcctgtagcgccgcggtttcagcatcacaggttattttg 180
               .        140          .        160          .        180

   ------------------------------------------------------------

   181 cctgaagcttgctggggcgtaaatccctctcgccttgtttctcagagagcatttcaggcc 240
               .        200          .        220          .        240

   ------------------------------------------------------------

   241 ggttttgcagtcgctgctgcagctatggggtccctagaaatggtgccaatgggcgcgggt 300
               .        260          .        280          .        300

   ------------------------------------------------------------

   301 ccccctagccccggcggggatccggatgggtacgatggcggaaacaactcccaatatcca 360
               .        320          .        340          .        360

   ------------------------------------------------------------

   361 tctgcttctggctcttctgggaacacccccaccccaccgaacgatgaggaacgtgaatct 420
               .        380          .        400          .        420

   ------------------------------------------------------------

   421 aatgaagagcccccaccgccttatgaggacccatattggggcaatggcgaccgtcactcg 480
               .        440          .        460          .        480

   ------------------------------------------------------------

   481 gactatcaaccactaggaacccaagatcaaagtctgtacttgggattgcaacacgacggg 540
               .        500          .        520          .        540

   ------------------------------------------------------------

   541 aatgacgggctccctccccctccctactctccacgggatgactcatctcaacacatatac 600
               .        560          .        580          .        600
```

```
     ------------------------------------------------------------

 601 gaagaagcgggcagaggaagtatgaatccagtatgcctgcctgtaattgttgcgccctac 660
               ·        620          ·        640          ·        660

     ------------------------------------------------------------

 661 ctcttttggctggcggctattgccgcctcgtgtttcacggcctcagttagtaccgttgtg 720
               ·        680          ·        700          ·        720

     ------------------------------------------------------------

 721 accgccaccggcttggccctctcacttctactcttggcagcagtggccagctcatatgcc 780
               ·        740          ·        760          ·        780

     ------------------------------------------------------------

 781 gctgcacaaaggaaactgctgacaccggtgacagtgcttactgcggttgtcactttcttt 840
               ·        800          ·        820          ·        840

     ------------------------------------------------------------

 841 gcaatttgcctaacatggaggattgaggacccacctttaattctcttctgtttgcattg 900
               ·        860          ·        880          ·        900

     ------------------------------------------------------------

 901 ctggccgcagctggcggactacaaggcatttacgttctggtgatgcttgtgctcctgata 960
               ·        920          ·        940          ·        960

     ------------------------------------------------------------

 961 ctagcgtacagaaggagatggcgccgtttgactgtttgtggcggcatcatgttttggca 1020
               ·        980          ·       1000          ·       1020

     ------------------------------------------------------------

1021 tgtgtacttgtcctcatcgtcgacgctgttttgcagctgagtcccctccttggagctgta 1080
               ·       1040          ·       1060          ·       1080

   1 -----------catg---------------------------------------------- 4
              ||||
1081 actgtggtttccatgacgctgctgctactggctttcgtcctctggctctcttcgccaggg 1140
               ·       1100          ·       1120          ·       1140

     ------------------------------------------------------------

1141 ggcctaggtactcttggtgcagccctttaacattggcagcagctctggcactgctagcg 1200
               ·       1160          ·       1180          ·       1200

     ------------------------------------------------------------

1201 tcactgattttgggcacacttaacttgactacaatgttccttctcatgctcctatggaca 1260
               ·       1220          ·       1240          ·       1260

     ------------------------------------------------------------

1261 cttgtggttctcctgatttgctcttcgtgctcttcatgtccactgagcaagatccttctg 1320
               ·       1280          ·       1300          ·       1320

     ------------------------------------------------------------

1321 gcacgactgttcctatatgctctcgcactcttgttgctagcctccgcgctaatcgctggt 1380
               ·       1340          ·       1360          ·       1380
```



```
              .       280         .       300         .       320
  263 gctcatgcacaccttggcgggtcatccgccacccccgtccagcaggctcaggccgccgca 322
       | | || ||| |||   |  |||  |  | | ||  |    ||  |||||||
 1978 ccccctggaca=ctt=atgtttcaagcagctcacctat=ggtca==ctcaggc======= 2025
    1980          .       2000          .       2020
              .       340         .       360         .       380
  323 tccgctgggaccggggccttggcatcatcagcgccgtccacggccgtagcccagtc-cgc 381
                                  |  | || || |  |    | ||||||  |
 2026 ==============================acggtcg=cccctccgagtgaccagtcacct 2055
                                       .       2040        .
              .       400                         .       420
  382 gacccccctctgtttcttcatctattagc----------------agcctccgggccgcga 425
        |   || ||  |   |     |||||                 ||||  |||||||
 2056 tccagactatgcatacactgaatttagcctgatattgtccccctagcc=ccgggcc==== 2110
       2060          .       2080          .       2100          .
          .       440         .       460         .       480
  426 cttcgggggcgactgccgccgcctccgccgccgcagccgtcgataccgggtcaggtggcg 485
                               |  | || || |||    |  | |  ||  ||  |  |
 2111 =======================cagc=cctcctcagaaaactctgcatgg==agaaagctg 2145
                                    2120          .       2140
          .       500         .       520         .
  486 ggggacaaccccacgacaccgccccacgcggggcacgtaagaaacagtagccc-------- 538
      |  |  |||| | |  | | | ||   | |  | |  | | |||||  ||
 2146 gacgtgaacctc=ccccccagacctgtgtgctgta=tttacaaacactacaataaaccca 2203
          .       2160          .       2180          .       2200

       -------------------------

 2204 atgtgcaaaaaaaaaaaaaaaaa                                    2227
              .       2220

 % Identity =   7.0 (174/2484)

///
```

DNA Strider 1.4f6 ### Friday, July 13, 2007 4:45:02 PM (US Letter @ 100%)

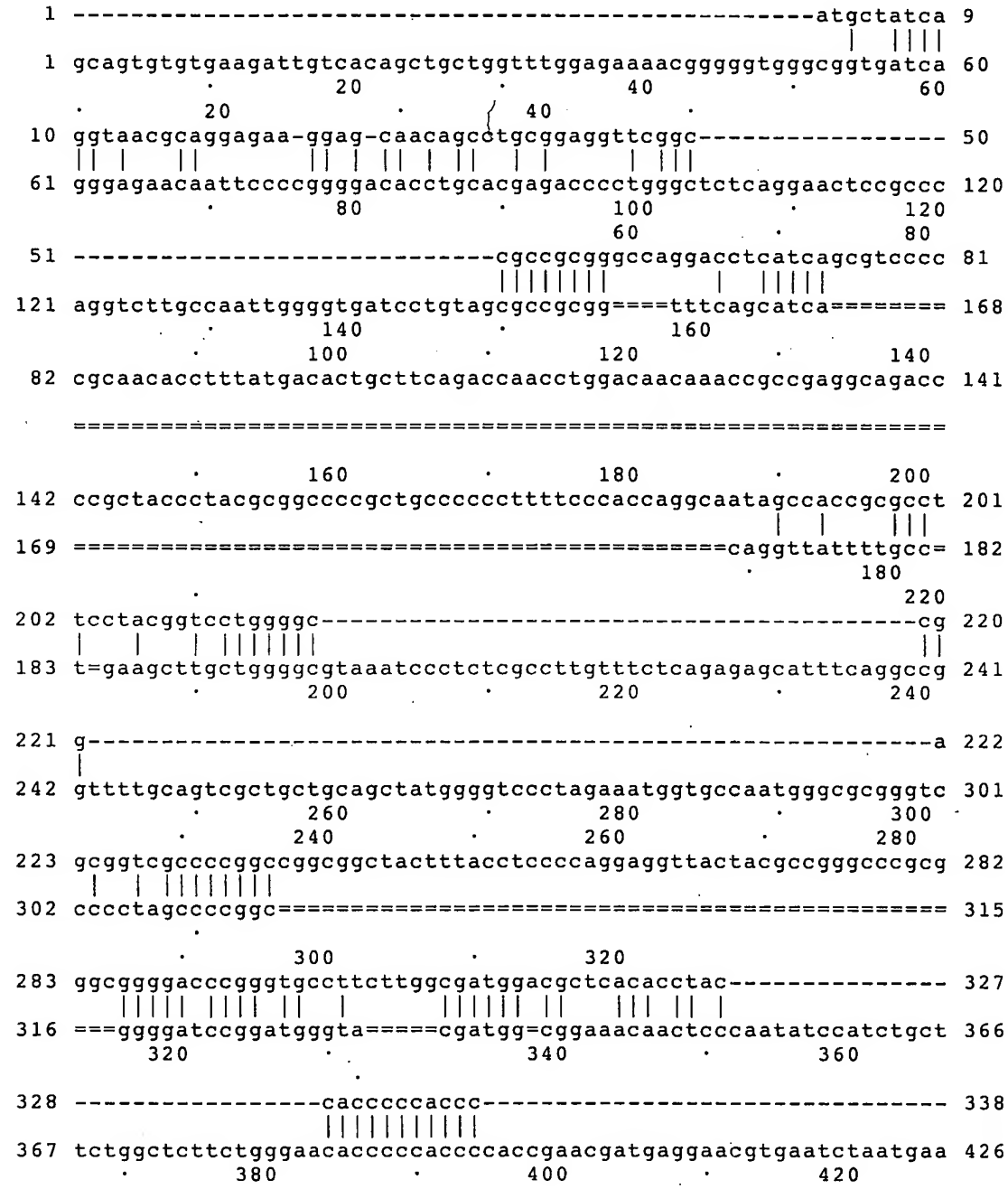
9310-13DVCTDV SEQ ID NO 3.xdna x Laux et al. EBV terminal gene.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

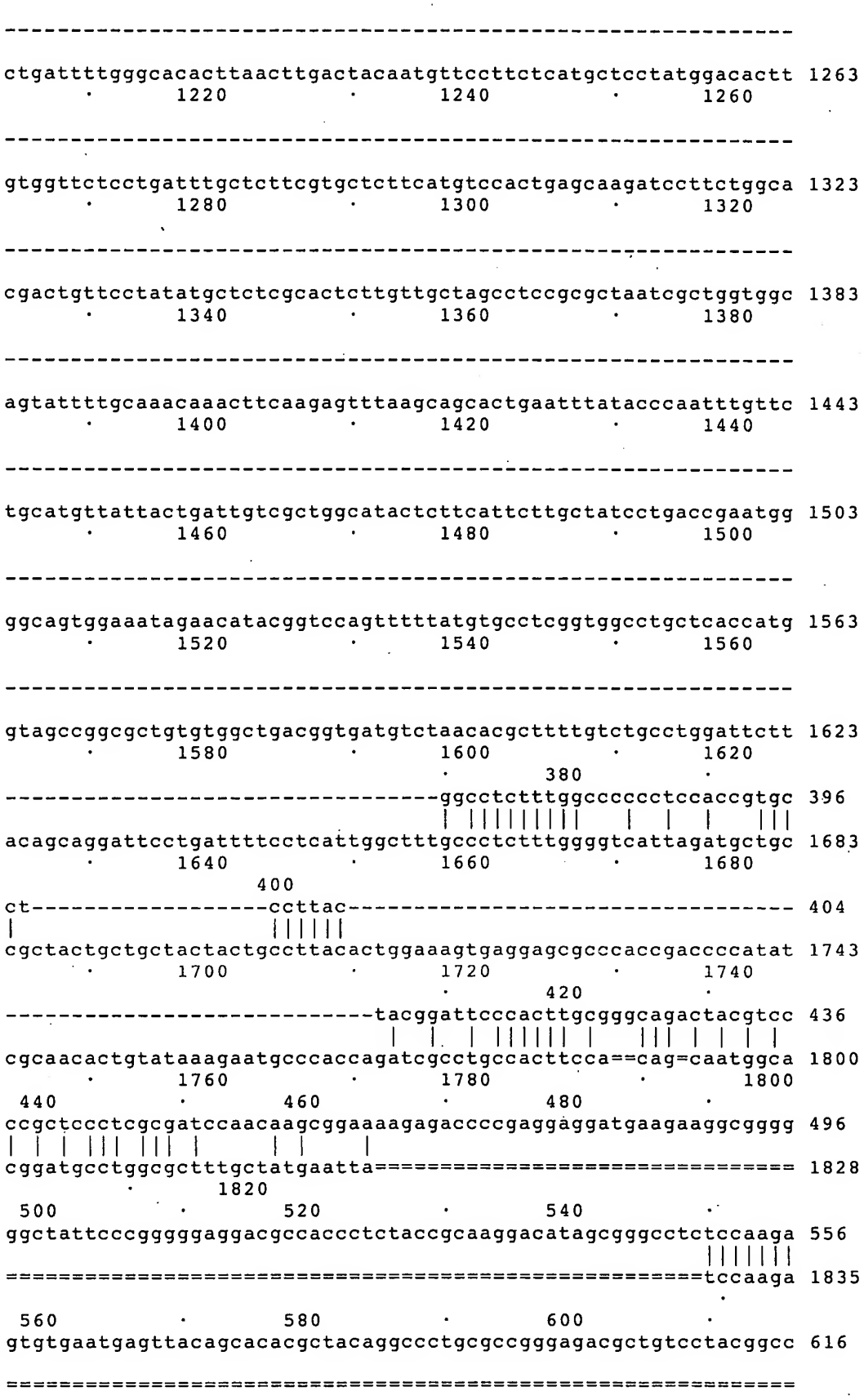
DNA sequence 2227 bp gcagtgtgtgaa ... aaaaaaaaaaaa linear

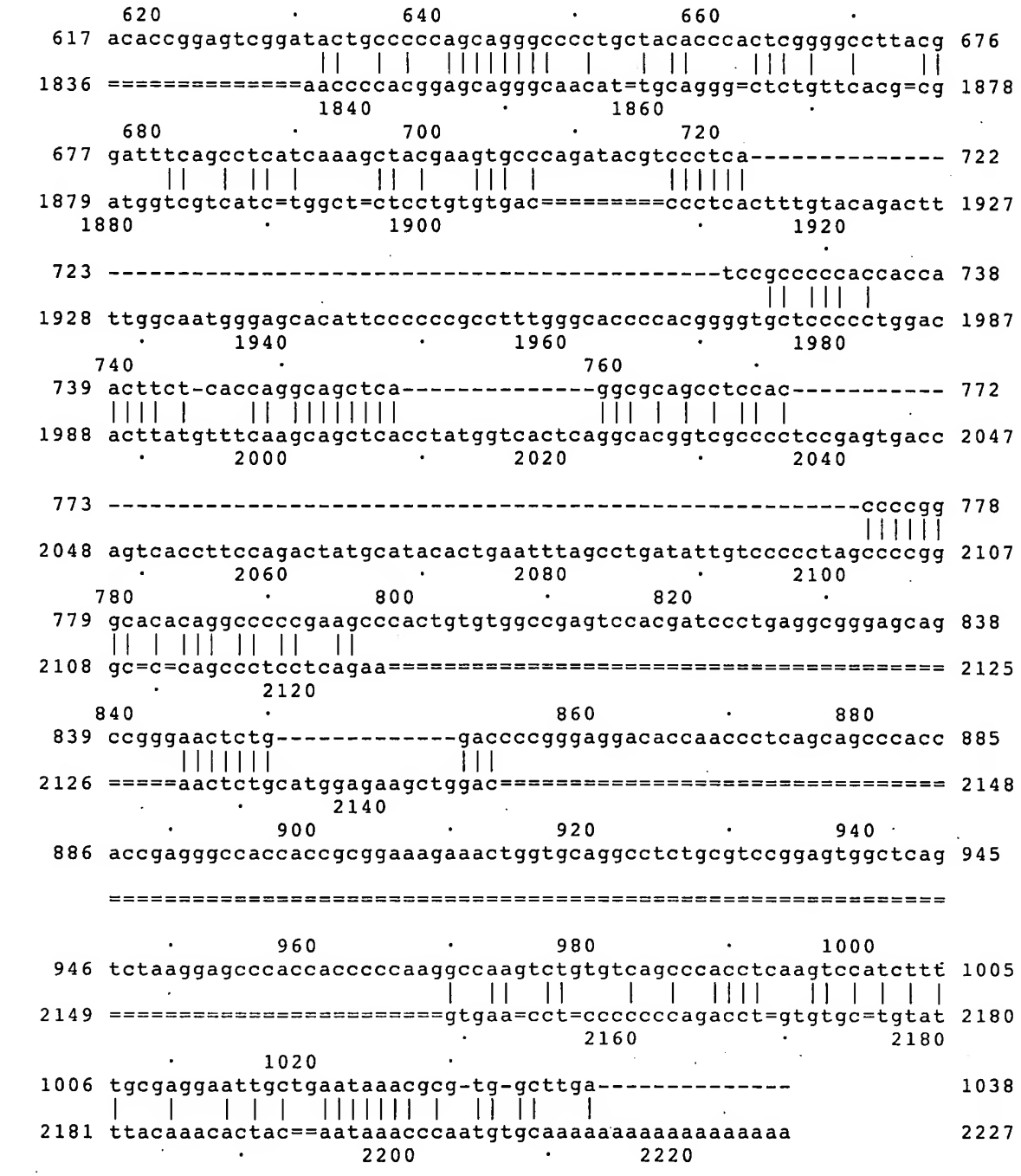
Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 2. Comparison of nucleotide sequence of SEQ ID NO:3 with the nucleotide sequence of Fig. 2 of Laux et al.



```
     ------------------------------------------------------------

 427 gagcccccaccgccttatgaggacccatattggggcaatggcgaccgtcactcggactat 486
         .       440         .       460         .       480

     ------------------------------------------------------------

 487 caaccactaggaacccaagatcaaagtctgtacttgggattgcaacacgacgggaatgac 546
         .       500         .       520         .       540
          340         .
 339 -----acacccccctccggcctac------------------------------------ 357
           |  ||||||||    ||||
 547 gggctccctcccccctcc===ctactctccacgggatgactcatctcaacacatatacgaa 603
         .       560         .       580         .       600

     ------------------------------------------------------------

 604 gaagcgggcagaggaagtatgaatccagtatgcctgcctgtaattgttgcgccctacctc 663
         .       620         .       640         .       660
      360
 358 -tttggct---------tgccg---------------------------------------- 369
      |||||||         |||||
 664 ttttggctggcggctattgccgcctcgtgtttcacggcctcagttagtaccgttgtgacc 723
         .       680         .       700         .       720

     ------------------------------------------------------------

 724 gccaccggcttggccctctcacttctactcttggcagcagtggccagctcatatgccgct 783
         .       740         .       760         .       780

     ------------------------------------------------------------

 784 gcacaaaggaaactgctgacaccggtgacagtgcttactgcggttgtcactttctttgca 843
         .       800         .       820         .       840

     ------------------------------------------------------------

 844 atttgcctaacatggaggattgaggacccacctttaattctcttctgtttgcattgctg 903
         .       860         .       880         .       900

     ------------------------------------------------------------

 904 gccgcagctggcggactacaaggcatttacgttctggtgatgcttgtgctcctgatacta 963
         .       920         .       940         .       960

     ------------------------------------------------------------

 964 gcgtacagaaggagatggcgccgtttgactgtttgtggcggcatcatgtttttggcatgt 1023
         .       980         .      1000         .      1020

     ------------------------------------------------------------

1024 gtacttgtcctcatcgtcgacgctgttttgcagctgagtcccctccttggagctgtaact 1083
         .      1040         .      1060         .      1080

     ------------------------------------------------------------

1084 gtggtttccatgacgctgctgctactggctttcgtcctctggctctcttcgccagggggc 1143
         .      1100         .      1120         .      1140

     ------------------------------------------------------------

1144 ctaggtactcttggtgcagccctttttaacattggcagcagctctggcactgctagcgtca 1203
         .      1160         .      1180         .      1200
```





% Identity = 11.7 (323/2749)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ*

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV*

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 3. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```
      ------------------------------------------------------------

    1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
                .         20          .         40          .         60

      ------------------------------------------------------------

   61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGSM 120
                .         80          .         100         .         120

      ------------------------------------------------------------

  121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
                .         140         .         160         .         180

      ------------------------------------------------------------

  181 PVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
                .         200         .         220         .         240

      ------------------------------------------------------------

  241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVSMTLLLLAFVLWLSSPGGLGTLGAA 300
                .         260         .         280         .         300
                                      .         20          .
    1 ----------------------MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPN 37
                               M    +    TL    L       S   L K    L +   L
  301 LLTLAAAALALLASLILGTLNLTTMFLLMLLWTLVVLLICSSCSSCPLSK=ILLARLFL=Y 358
                .         320         .         340         .
       40         .         60         .         80         .
   38 DVFREAQRSYLVFLTSQFCYEEYVQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSAT 97
       +       S L+    S           +         +      A +     A      GS
  359 ALALLLLASALIAGGSILQTNFKSLSSTEFIPNLFCMLLLIVAGILFILAILTEWGSGNR 418
     360         .         380         .         400         .
      100        .         120         .         140         .
   98 PVQQAQAAASAGTGALASSAPSTAVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTG 157
                      +A +    T ++ +    +    +    L       A
  419 TYGPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCY 478
     420         .         440         .         460         .
      160        .
  158 SGGGGQPHDTAPRGARKKQ*                                        177
            +   +   P    R    *
  479 YCLTLESEERPPTPYRNTV*                                        498
     480         .
```

% Identity = 4.6 (23/500) % Homology = 3.2 (16/500) % Total = 7.8 (39/500)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA*

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV*

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 4. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```
                 .         20         .         40         .
    1 MLSGNAGE-GATACG-GSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPF 58
      M S      GA     G    G D  +  +    +    N    P  +        P PP+
    1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
              .         20         .         40         .         60
      60          .         80         .        100         .
   59 SHQAIATAP--S-YGP-GAGAVAPAGGYFTSPG-GYYAGPAGG-DPGAFLAMD-AHTYHP 111
                 S Y P G    +   G       G    P    D  +    + A
   61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGSM 120
              .         80         .        100         .         120
            120          .        140         .        160
  112 HPHPPPAYFG--LPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLF- 168
      +P   P      L  L      C   + S +      A S                  L
  121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
              .        140         .        160         .         180
       .        180         .        200         .        220
  169 PGEDATLYRKDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQ 228
      P    T      A       E       L     + G     Y          +    ++
  181 PVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
              .        200         .        220         .         240
       .        240         .        260         .        280
  229 PHQSYEVPRYVPHPPPPPTSHQAAQAQPPPPGTQAPEAHCV-AESTIPEA-GAAGNSGPR 286
              ++               +      T       + A      + G  G  G
  241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVSMTLLLLAFVLWLSSPGGLGTLGAA 300
              .        260         .        280         .         300
         .       300         .        320         .        340
  287 EDTNPQQPTTEGHHRGKKLVQASASGVAQSKEPTTPKAKSV-SAHLKS-IFCEELLNKRV 344
        T               L   +   +            S  S+   S I    L    +
  301 LLTLAAALALLASLILGTLNLTTMFLLMLLWTLVVLLICSSCSSCPLSKILLARLFLYAL 360
              .        320         .        340         .         360

  345 A*------------------------------------------------------------ 346
      A
  361 ALLLLASALIAGGSILQTNFKSLSSTEFIPNLFCMLLLIVAGILFILAILTEWGSGNRTY 420
              .        380         .        400         .         420

      ------------------------------------------------------------

  421 GPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
              .        440         .        460         .         480

      ------------------

  481 LTLESEERPPTPYRNTV*                                          498
              .
```

% Identity = 10.0 (50/498) % Homology = 4.8 (24/498) % Total = 14.9 (74/498)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

```
Protein  sequence       24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence      498  aa   MGSLEMVPMGAG ... ERPPTPYRNTV*

   Method:           Diagonals (BLOSUM62)
   Layout:           Standard
   Block Length =:   6-aa
   Mismatch penalty: Smaller (1)
   Gap penalty:      Medium  (2)
   Display:          BLOSUM62
```

Alignment 5. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```
                    .         20
     1 AVDTGSGG-G-GQPHDTA-PRGARKKQ---------------------------------- 24
                 G G P     P G
     1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
                 .         20          .         40          .         60

       ------------------------------------------------------------

    61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGSM 120
                 .         80          .        100          .        120

       ------------------------------------------------------------

   121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
                 .        140          .        160          .        180

       ------------------------------------------------------------

   181 PVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
                 .        200          .        220          .        240

       ------------------------------------------------------------

   241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVSMTLLLLAFVLWLSSPGGLGTLGAA 300
                 .        260          .        280          .        300

       ------------------------------------------------------------

   301 LLTLAAALALLASLILGTLNLTTMFLLMLLWTLVVLLICSSCSSCPLSKILLARLFLYAL 360
                 .        320          .        340          .        360

       ------------------------------------------------------------

   361 ALLLLASALIAGGSILQTNFKSLSSTEFIPNLFCMLLLIVAGILFILAILTEWGSGNRTY 420
                 .        380          .        400          .        420

       ------------------------------------------------------------

   421 GPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
                 .        440          .        460          .        480

       ------------------

   481 LTLESEERPPTPYRNTV*                                          498
                 .

  % Identity =   1.0 (5/498)       % Homology =   0.0 (0/498)       % Total =   1.0 (5/498)

///
```

9310-13DVCTDV SEQ ID NO 6.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

```
Protein  sequence       30  aa   STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence      498  aa   MGSLEMVPMGAG ... ERPPTPYRNTV*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 6. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```
      ------------------------------------------------------------

    1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
                .         20          .         40          .         60

      ------------------------------------------------------------

   61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGSM 120
                .         80          .        100          .        120

      ------------------------------------------------------------

  121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
                .        140          .        160          .        180

      ------------------------------------------------------------

  181 PVTVLTAVVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
                .        200          .        220          .        240

      ------------------------------------------------------------

  241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGAVTVVSMTLLLLAFVLWLSSPGGLGTLGAA 300
                .        260          .        280          .        300

      ------------------------------------------------------------

  301 LLTLAAALALLASLILGTLNLTTMFLLMLLWTLVVLLICSSCSSCPLSKILLARLFLYAL 360
                .        320          .        340          .        360

      ------------------------------------------------------------

  361 ALLLLASALIAGGSILQTNFKSLSSTEFIPNLFCMLLLIVAGILFILAILTEWGSGNRTY 420
                .        380          .        400          .        420
                                                                      .
    1 ------------------------------------------------STAVAQSATPSV 12
                                                                  +
  421 GPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
                .        440          .        460          .        480
          20          .
   13 SSSISSLRAATSGATAAA                                          30
       +   S   R   T
  481 LTLESEERPPTPYRNTV*                                          498
                .

  % Identity =   0.6 (3/498)       % Homology =   0.4 (2/498)       % Total =   1.0 (5/498)

///
```

9310-13DVCTDV SEQ ID NO 1.xdna x Bankier et al. EcoRI Dhet fragment.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

DNA sequence 12436 bp gaattctcaaag ... tgtttagaattc linear

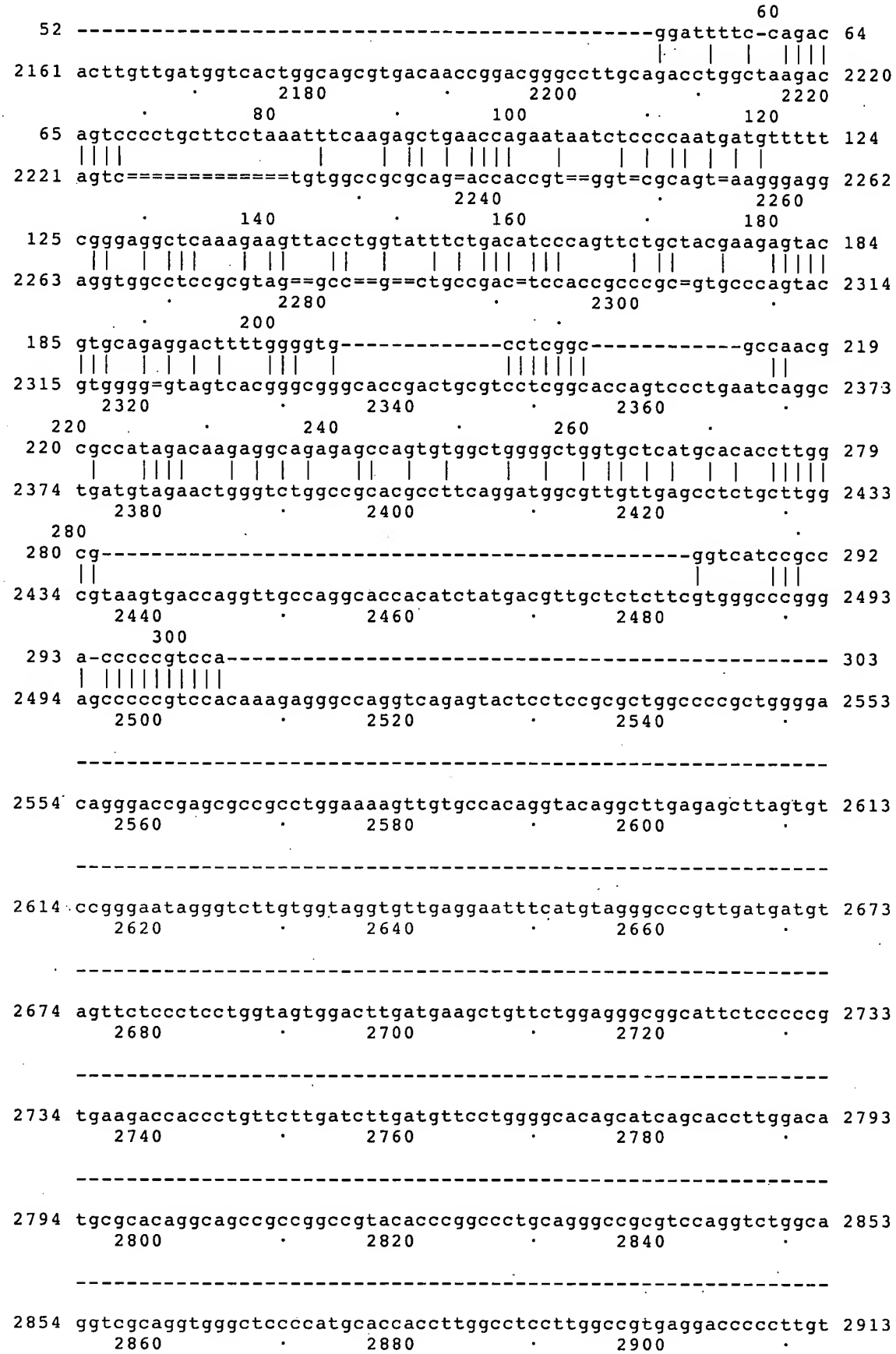
Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 7. Comparison of nucleotide sequence of SEQ ID NO:1 with the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 gaattctcaaaggcggcaccctcgccggcgcgcctgtcctcccagggacccgagacgaag 60
                ·         20         ·         40         ·         60

      ------------------------------------------------------------

   61 gcccgtctgtagaggaagtggttgcgcatgcgggccagctcccagtagaccacgtccccc 120
                ·         80         ·        100         ·        120

      ------------------------------------------------------------

  121 cagacgcgcaggcacagggtctcggtcagggtctcgctctgttgcgccaggcaggactgc 180
                ·        140         ·        160         ·        180

      ------------------------------------------------------------

  181 agcttggccagaccctcggtggccacctggcgcaggtactgctccttgcgcttgagcgcg 240
                ·        200         ·        220         ·        240

      ------------------------------------------------------------

  241 tccgagagggcgccggacgggccgggctctcgtgccccagccggccggggcacctccggg 300
                ·        260         ·        280         ·        300

      ------------------------------------------------------------

  301 ctctcccgggacgcctcctcctcgcctcggcccaaccgctgcatggctcggttgagccgc 360
                ·        320         ·        340         ·        360

      ------------------------------------------------------------

  361 gtgtacagctcgttcctcttttgcaggatggcccggtactgggggtgcgccgtgaaggcg 420
                ·        380         ·        400         ·        420

      ------------------------------------------------------------

  421 gcggcgcagtccgccttcagcgcctccaccgcgtcgcccgaggagctgtagaccccgccg 480
                ·        440         ·        460         ·        480

      ------------------------------------------------------------

  481 cagaagagccgctccgtggccccgggagccacggcgtcaaacaggtgagtcagccttgcc 540
                ·        500         ·        520         ·        540

    1 ------------------------------------------------------------catg 4
                                                                    |
  541 cccgccagcgcctcctcgcaggcccccgcaccagggccaggcgacgctcccgggcaaac 600
                ·        560         ·        580         ·        600
```

```
                          20
   5 atggcacgccggctgcccaag-c-ccaccctcc-------------------------- 35
     | ||||    ||| |      | | |||||||||
 601 agggcagagaggcgggaatggccgccaccctcccccctgccccgttgcaccgatagcatgg 660
                 620          640          660

     ------------------------------------------------------------
 661 ccgccagagttccaatagaggagctccgagagctccgccacctccgggggcactgtcgag 720
                 680          700          720

     ------------------------------------------------------------
 721 aagacgttgtaggtgtccagcgctctggtcgccccctctgcctccggccgccccgggccc 780
                 740          760          780

     ------------------------------------------------------------
 781 gggaccgcgccctcctctgggccgcccggcctcgccttctcctcagcctccaacaggtgc 840
                 800          820          840

     ------------------------------------------------------------
 841 ccgagccccgcctggcggacttcattctcaaacagtcccgagaccggctccggattcacc 900
                 860          880          900

     ------------------------------------------------------------
 901 ggcaccgccaggtggttacaggagacgtgggtcccctctgccgtggaagggttgccgtgg 960
                 920          940          960

     ------------------------------------------------------------
 961 ttgggcagaaccatcagctcgcccacacagcgccagcagggcacagaggtgatgtagagg 1020
                 980          1000         1020

     ------------------------------------------------------------
1021 cgcgggtctgggatgggacttacgccccgaaagcggcccagcagatccagggcccgttcc 1080
                 1040         1060         1080

     ------------------------------------------------------------
1081 aggctctccagccccatggtgtgagacatgcaataaaacacgctattgattctcttcatt 1140
                 1100         1120         1140

     ------------------------------------------------------------
1141 aaaatctctatgtcatttattaggcacaaacttacatcgactttatgccccccgtaaaac 1200
                 1160         1180         1200

  36 -----------------------------------------------------------a 36
                                                                |
1201 tccacagagtacgcgactgaggggtacggagaggcgggacccgggtaccctttctacca 1260
                 1220         1240         1260
      40
  37 ggggag------------------------------------------------------ 42
     |||| |
1261 ggggcgagcagcgcggcagaggcctctctcgagttctctagcaggtgcaccagctccagg 1320
                 1280         1300         1320

     ------------------------------------------------------------
1321 gacagggcgctgcatgcacggtcattctgccgtctcaaacggggaaggaggatggcctcc 1380
                 1340         1360         1380
```

```
                                                                        
1381 agctcggccagcaggccggcgttgcgcaccaccgcagccacgtccagactccgggggtcc 1440
                1400          1420          1440                      

                                                                        
1441 agccgggtgcacacgctcagctcaaccgccagggcgtacacctggctgtacgccgccgcc 1500
                1460          1480          1500                      

                                                                        
1501 agcagccccgacatcgccgccccaggggtctctagacctcgagtccggggagaacggtgg 1560
                1520          1540          1560                      

                                                                        
1561 ccagacggcgcttgcgtctgcccccggagccctgccctcctccacccagcagcagcccgg 1620
                1580          1600          1620                      

                                                                        
1621 ccgaggcctgcgacgcggtgctgaccggctcggccacgctgataaagttgtcctgggctg 1680
                1640          1660          1680                      

                                                                        
1681 ccccgggcccaccccacactccctccagaaagtcccgagcggcctccgccgtccactcta 1740
                1700          1720          1740                      

  43 ----gctggaggc---------------------------------------------------- 51
         |||||||||
1741 tcccgctggaggcaatggtcgccagggtttctaggacgctgtccgccaggacggagaagc 1800
                1760          1780          1800                      

                                                                        
1801 ggcccaataagtactccgcgtcgtccctagtcagcgaggcgcatgcctcgcccatggcat 1860
                1820          1840          1860                      

                                                                        
1861 ccacaaggttgcacaccacatcaaacacacagtcttcctcctgtttttgtgatataatgg 1920
                1880          1900          1920                      

                                                                        
1921 cctccaggccagccctgatgttctcaatctcatatgtggtcgcggcttgggtccggcgct 1980
                1940          1960          1980                      

                                                                        
1981 tcacggtcaaccctagggtgggggtggcaaagacaaacttcttccgcatggaagagcccc 2040
                2000          2020          2040                      

                                                                        
2041 cggcctgcttgcgcagcccagccccgggggcctgcagcaggttcctgtccacgccccggc 2100
                2060          2080          2100                      

                                                                        
2101 ccataaagtatcccaggttcccggcctggaatatctggttgttgccgttgacccccgtgt 2160
                2120          2140          2160                      
```



```
 304 ------gcaggctc------------------------------------------------ 311
           |||||||
2914 cgatggccaggctcctaaagttggtgcacagcgtctggtagtgacccttagccactctg 2973
     2920          2940          2960

     ------------------------------------------------------------

2974 gggggctctggccaagcccggggttgtcattctcatagcacatacagatgggcagggaga 3033
     2980          3000          3020

     ------------------------------------------------------------

3034 tgtcctgcaggatggtcagcagtgagcggtaaaacagctgggtgaagatggggcaggcgg 3093
     3040          3060          3080

     ------------------------------------------------------------

3094 gctgcgcaaaggggttgcacgagtactgcatcacgtggtagcagctcttgaccaggtcct 3153
     3100          3120          3140

 312 ----------------------------------------------------------aggccg 317
                                                                |||
3154 tgtaggtgatgttgttcttggccatgctgttcataaactggaccacttcggcgtccaccg 3213
     3160          3180          3200
     320
 318 ccgcatcc------------------------------------------------------ 325
     ||||||||
3214 ccgcatccacgtccttgaacatcttgacaaagtcacgcgggccatggggctccttctcta 3273
     3220          3240          3260

     ------------------------------------------------------------

3274 gctttccttcagcgtctatgcccagccgagacagccgctccagcaggttctggttcagct 3333
     3280          3300          3320

     ------------------------------------------------------------

3334 gccagtaggtgtagcggggctcgtcgtccggccgctgcccgtcgtcctccttatcgatga 3393
     3340          3360          3380

     ------------------------------------------------------------

3394 agttgagaaagttgcccaaaaagtccgtctcgttgtaggagcccgaggcccccgagatca 3453
     3400          3420          3440

 326 ----------------gctgg------------------------------------------ 330
                     ||||x
3454 catàggggtccctccgctgcgtggacatgacgggggggaagcggtccctcagcctaaaga 3513
     3460          3480          3500

     ------------------------------------------------------------

3514 agagcgtgttcaggcacacggccggggcccggccctcgcagagcgagcacatgggactgg 3573
     3520          3540          3560

     ------------------------------------------------------------

3574 cggccgcccccgccacgtagctgcccgtctccggcaccggggtcagagagctcttctgtc 3633
     3580          3600          3620

     ------------------------------------------------------------

3634 cctggcaaaactgcaggtagtaggcatagcgggcaagaaggttgggcgagaaggaggccg 3693
     3640          3660          3680
```

```
     ------------------------------------------------------------

3694 catagaccaggtgctccacagcgtagtttcccggaccgttggttccggtcacgtctggcc 3753
        3700          .       3720          .       3740          .

     ------------------------------------------------------------

3754 caccccagcccgagaagcagggtcggcggcaggggtcccaggtcccctcctgcagggtcc 3813
        3760          .       3780          .       3800          .

     ------------------------------------------------------------

3814 ccaggccgtgggtcatgtagaaactgttaaagagactctccttgccctgaccggttgact 3873
        3820          .       3840          .       3860          .

     ------------------------------------------------------------

3874 tcgagacccccgagacgtagaggacggaattggtggcaaagatctgcgtggacacgtggg 3933
        3880          .       3900          .       3920          .

     ------------------------------------------------------------

3934 gggccaggctggcattatatcggtgtaacgcagccacacgggcctctggaccctcacagt 3993
        3940          .       3960          .       3980          .

     ------------------------------------------------------------

3994 cggcaaacaggggccacgagtcgtagttgaggctggccggggtctcgtgcgaggcctcca 4053
        4000          .       4020          .       4040          .

     ------------------------------------------------------------

4054 gcatggcgggtgcgtagctcaccgccagctcgcatgccgcgctgtccacaatcattaagg 4113
        4060          .       4080          .       4100          .

     ------------------------------------------------------------

4114 ctcccgagtccgggtgactgatggttgaggctgggaactccttgaggggggccaccttgg 4173
        4120          .       4140          .       4160          .

     ------------------------------------------------------------

4174 ccaccttggcctggtcctgcaggctctgcttctccagcagctccaccagcttgcccaccc 4233
        4180          .       4200          .       4220          .

     ------------------------------------------------------------

4234 gtcggacgcgcagcgcctgcgccagcccggtgtacagcgcctcgtgcatgcagcggctga 4293
        4240          .       4260          .       4280          .

     ------------------------------------------------------------

4294 ggtccgagttgtaaaactggcggagctggggcacgccctctgggaacacctccttgtcgt 4353
        4300          .       4320          .       4340          .

     ------------------------------------------------------------

4354 agagcgggaccctaacgctcgcagactgccccaccgctacctcctgttttaacgatggaa 4413
        4360          .       4380          .       4400          .

     ------------------------------------------------------------

4414 tggccaccaggtttccgctgtagagtcgctccttgaaggcctcggttattgccaccgccc 4473
        4420          .       4440          .       4460          .
```

```
4474 ccaggtaggcagagggatctagcccttcggggaagaagtcccccggcttggagctttccc 4533
        4480          .       4500          .       4520          .

4534 tcggtagggcgctgtaggcgtcgtacccaaacacctccctggtctcgccacagagggcct 4593
        4540          .       4560          .       4580          .

4594 cgagacccggcccctcaaagatgggggggaaccatatgggcattgtggaacacgtagatgt 4653
        4600          .       4620          .       4640          .

4654 ccctgtgataggaggtagcgcgtaggagcccgcagttggggtcgggcctcctgtgcagag 4713
        4660          .       4680          .       4700          .

4714 ccttgacattgatgctgaagcccggctccacggtgatgccgcaaaggagcggcaccgtca 4773
        4720          .       4740          .       4760          .

4774 ggcacctgtggcccgcgtagccggtccccagtgtggccacctccctaagagggtaggtgg 4833
        4780          .       4800          .       4820          .

4834 ccaggggggtaaaagtagatgtagccgcacggacccggctggctctggctgcccagattat 4893
        4840          .       4860          .       4880          .

4894 cctcgctagtctgtgcaccctgcatgatgcccaaggtatcgccccggcctcccagtccca 4953
        4900          .       4920          .       4940          .

4954 cattaaatgttacactttactcatcacgcaacacccactgtttattcatttacaaagatt 5013
        4960          .       4980          .       5000          .

5014 tcaggaagtcagtcaggctggccagggcccacgtcacggggaactgacgtctcagcgatc 5073
        5020          .       5040          .       5060          .

5074 ttggcatgccgcccagcctcgcaaaccagagtctgcgatagagggccaggtagtgggcga 5133
        5080          .       5100          .       5120          .

5134 ttgcccccagcacgaaggcggcgctcttgtggtcatccaggtagtttcgcaccgcaaaca 5193
        5140          .       5160          .       5180          .

5194 ccactgtgtagcacagcaccaccctgagccgcgaccagtagtcgtagtggtcgttgtaca 5253
        5200          .       5220          .       5240          .
```

```
     ------------------------------------------------------------

5254 ctgcgcgcaggacgctgatgatgagccgtacgtgcgtgtctttgcccccgatgtcggctg 5313
        5260          .       5280          .       5300          .

     ------------------------------------------------------------

5314 tcctgcaggccagctccgcgtacagcttcctatccttcctcagggaggccttgatgagcc 5373
        5320          .       5340          .       5360          .

     ------------------------------------------------------------

5374 ggcagaggaccagggctggcaaaggcaggtctttctcatcccgggtgaacaccgcgtaca 5433
        5380          .       5400          .       5420          .

     ------------------------------------------------------------

5434 tggccctgaacatgaggtagctggactcagccaccttgtcgtccggcggcgagggcgcga 5493
        5440          .       5460          .       5480          .

 331 ----------gaccgggg---------------------------------------------- 338
             ||||||||
5494 cccacgcctcgaccggggtcctcacaaacacagaatctgtagacttggctggcctcatgg 5553
        5500          .       5520          .       5540          .

     ------------------------------------------------------------

5554 tctcgtcaggccagctcacgggcttcaggcttatatgataaaatgggcgtggcagaatag 5613
        5560          .       5580          .       5600          .

     ------------------------------------------------------------

5614 tataagacgcgaggcctgggtgaggagagtccagagcaatggccaggttcatcgctcagc 5673
        5620          .       5640          .       5660          .

     ------------------------------------------------------------

5674 tcctcctgttggcctcctgtgtggccgccggccaggctgtcaccgctttcttgggtgagc 5733
        5680          .       5700          .       5720          .

     ------------------------------------------------------------

5734 gagtcaccctgacctcctactggaggagggtgagcctcggtccagagattgaggtcagct 5793
        5740          .       5760          .       5780          .

     ------------------------------------------------------------

5794 ggtttaaactgggcccaggagaggagcaggtgcttattgggcgcatgcaccacgatgtca 5853
        5800          .       5820          .       5840          .

     ------------------------------------------------------------

5854 tctttatagagtggcctttcaggggcttctttgatatccacagaagtgccaacaccttct 5913
        5860          .       5880          .       5900          .

     ------------------------------------------------------------

5914 ttttagtagtcaccgctgccaacatctcccatgacggcaactacctgtgccgcatgaaac 5973
        5920          .       5940          .       5960          .

     ------------------------------------------------------------

5974 tgggcgagaccgaggtcaccaagcaggaacacctgagcgtggtgaagcctctaacgctgt 6033
        5980          .       6000          .       6020          .
```

```
     ------------------------------------------------------------

6034 ctgtccactccgaaaggtctcagttcccagacttctctgtccttactgtgacatgcaccg 6093
        6040          .       6060          .       6080          .

     ------------------------------------------------------------

6094 tgaatgcatttccccatccccacgtccagtggctcatgcccgagggcgtggagcccgcac 6153
        6100          .       6120          .       6140          .

     ------------------------------------------------------------

6154 caactgcggcaaatggcggtgttatgaaggaaaaggatgggagcctctctgttgctgttg 6213
        6160          .       6180          .       6200          .

     ------------------------------------------------------------

6214 acctgtcacttcccaagccctggcacctgccagtgacctgcgttgggaaaaatgacaagg 6273
        6220          .       6240          .       6260          .

     ------------------------------------------------------------

6274 aggaagcccacggggtttatgtttctggatacttgtcgcaataaacgcacttgcctattt 6333
        6280          .       6300          .       6320          .

     ------------------------------------------------------------

6334 caccttgttttagtgtggcattgggggggtggcattgcgggtggatagcctcgcgactcg 6393
        6340          .       6360          .       6380          .

     ------------------------------------------------------------

6394 tgggaaaatgggcggaagggcaccgtgggaaaatagttccaggtgacagcagcagtgtgt 6453
        6400          .       6420          .       6440          .

     ------------------------------------------------------------

6454 gaagattgtcacagctgctggtttggagaaaacgggggtgggcggtgatcagggagaaca 6513
        6460          .       6480          .       6500          .

     ------------------------------------------------------------

6514 attccccggggacacctgcacgagacccctgggctctcaggaactccgcccaggtcttgc 6573
        6520          .       6540          .       6560          .

     ------------------------------------------------------------

6574 caattggggtgatcctgtagcgccgcggtttcagcatcacaggttattttgcctgaagct 6633
        6580          .       6600          .       6620          .

     ------------------------------------------------------------

6634 tgctggggcgtaaatccctctcgccttgtttctcagagagcatttcaggccggttttgca 6693
        6640          .       6660          .       6680          .

     ------------------------------------------------------------

6694 gtcgctgctgcagctatggggtccctagaaatggtgccaatgggcgcgggtccccctagc 6753
        6700          .       6720          .       6740          .

     ------------------------------------------------------------

6754 cccggcggggatccggatgggtacgatggcggaaacaactcccaatatccatctgcttct 6813
        6760          .       6780          .       6800          .
```

```
6814 ggctcttctgggaacacccccaccccaccgaacgatgaggaacgtgaatctaatgaagag 6873
        6820          .       6840          .       6860          .

6874 cccccaccgccttatgaggacccatattggggcaatggcgaccgtcactcggactatcaa 6933
        6880          .       6900          .       6920          .

6934 ccactaggaacccaagatcaaagtctgtacttgggattgcaacacgacgggaatgacggg 6993
        6940          .       6960          .       6980          .

6994 ctccctcccccctccctactctccacgggatgactcatctcaacacatatacgaagaagcg 7053
        7000          .       7020          .       7040          .

7054 ggcagaggaaggtaagagtgccatctatctgtacttttatttattgcatcacaagtcaca 7113
        7060          .       7080          .       7100          .

7114 tcaataataagggcgccatctagcgggagatgttatccacaccatcccaattcacatctc 7173
        7120          .       7140          .       7160          .

7174 agggacaacaggtcaaagttctttgttgacacccccagcgctggctccaggggggtggaag 7233
        7180          .       7200          .       7220          .

7234 cgttggatgcagtcctccgcatcggggcggacgcctcctcccaacgcgtttctgcggatc 7293
        7240          .       7260          .       7280          .

7294 agtcgctggctggtgggcatcggagtcggtgggcggtcctccacggggacacgctccttc 7353
        7300          .       7320          .       7340          .

7354 ttggccttgttctttgacctttttggacattcttctgaaggaacggcggagagtagcgtag 7413
        7360          .       7380          .       7400          .

7414 aatccagccagtggtctacccggtcgcatggtggcttcttagatgaggagcaggcataaa 7473
        7420          .       7440          .       7460          .

7474 agtccaaacaggacacagagtaccaccaggagtagtcttagtctgctgacgtctgggtcc 7533
        7480          .       7500          .       7520          .

7534 tcggggcaggggtggctaggcctggtctccgtagaagagccgggcaggccgcaggcagag 7593
        7540          .       7560          .       7580          .
```

```
      ------------------------------------------------------------

7594  gactgctgctctagcaaagcacgctccaggacgtgtaccatctcgagagtgaggcacagc 7653
         7600          .       7620          .       7640          .

      ------------------------------------------------------------

7654  tgttttcgtggacttttatacagtaaggacaaggaaagaaggccagaggaatgtggaaag 7713
         7660          .       7680          .       7700          .

      ------------------------------------------------------------

7714  atgagcgaggacaggtgtggaggttttgggctagctcttagtttctgggtgtgagagagg 7773
         7720          .       7740          .       7760          .

      ------------------------------------------------------------

7774  gattaaagtgcttatgcgcaaagaatgtgtcaacaacaggtgttcctgcctctgctggca 7833
         7780          .       7800          .       7820          .

      ------------------------------------------------------------

7834  tgagttaggtgtggcttgggctgaatccaaatgtgtattggcacaagatggaaagcaaag 7893
         7840          .       7860          .       7880          .

      ------------------------------------------------------------

7894  ttgctggagttactgggtgggagacagggatgtatgtggtcccccgctggtatgccagta 7953
         7900          .       7920          .       7940          .
                     340
 339  -------------------cct------------------------------------- 341
                          |||
7954  ccctgtggaagtaaggggcctcatctgcctggtagttgtgttgtgcagaggtctgatgtg 8013
         7960          .       7980          .       8000          .

      ------------------------------------------------------------

8014  tgtaggaggggtgggttcaacgcaggggcgttggtggcggagtctggcaacgcccgggtc 8073
         8020          .       8040          .       8060          .

      ------------------------------------------------------------

8074  cttgctacctgtgtggtgtgttaagggctgggtaaaggtgtctgccaattctcgcatgtc 8133
         8080          .       8100          .       8120          .

      ------------------------------------------------------------

8134  ctcctttccccttgttttgaaatagaatatgaatgtggcttttcagcctagacagacagt 8193
         8140          .       8160          .       8180          .

      ------------------------------------------------------------

8194  gtggctaagggagtgtgtgccagttaaggtgattagctaaggcattcccagtaaatggag 8253
         8200          .       8220          .       8240          .

      ------------------------------------------------------------

8254  ggagagtcagtcaggcaagcctatgacatggtaatgcctagaagtaaagaaaggttagtc 8313
         8260          .       8280          .       8300          .
                                                .       360
 342  ---------------------------------tggcatcatcagcgccgtcca-cggccg 368
                                       | |||||||       |||| |||
8314  atagtagcttagctgaactgggccgtgggggtcgtcatcatc=======tccaccggaac 8366
         8320          .       8340          .               8360
```

```
      ·       380         ·       400         ·       420
 369 tagcccagtccgcgacccccctctgtttcttcatctattagcagcctccgggccgcgactt 428
     ||    |  ||   | |  |   |     |  |  || | | ||| ||  | | ||   |
8367 cagaagaacccaaaagcagcgtaggaaggt=gtggatca=ccgccgccatggc=cggaat 8423
        ·      8380         ·      8400         ·      8420
      ·       440         ·       460
 429 cgggggcgactgccgccgcctccgcc-gccgcagccgtc---------------------- 466
     |   | | |   |||||||||||| | | |       | |
8424 c=atgactatgaccgccgcctccgtctgtcatcaaaggcgggccctggtcacctcctttg 8482
           ·      8440         ·      8460         ·      8480

     ------------------------------------------------------------

8483 ttttcaacctcttccgtcaattgtggagggcctccatcatttccagcagagtcgctaggg 8542
           ·      8500         ·      8520         ·      8540

     ------------------------------------------------------------

8543 ctatgaggcagcgggtcatgtgggccattgtcatcagtgttgtcagggtcctgtgggcca 8602
           ·      8560         ·      8580         ·      8600

     ------------------------------------------------------------

8603 ttgtcatcagtgttgtcagggtcctgaggcagcgggtcatgtgggccattgtcatcagtg 8662
           ·      8620         ·      8640         ·      8660

     ------------------------------------------------------------

8663 ttgtcagggtcctgtgggccattgtcatcagtgttgtcagggtcctgtgggccattgtca 8722
           ·      8680         ·      8700         ·      8720

     ------------------------------------------------------------

8723 ggaccacctccaggtgcgcctaggttttgagagcagagtgggggtccgtcgccggctcca 8782
           ·      8740         ·      8760         ·      8780

     ------------------------------------------------------------

8783 ctcacgagcaggtggtgtctgccctcgttggagttagagtcagattcatggccagaatca 8842
           ·      8800         ·      8820         ·      8840

     ------------------------------------------------------------

8843 tcggtagcttgttgagggtgcgggagggagtcatcgtggtggtgttcatcactgtgtcgt 8902
           ·      8860         ·      8880         ·      8900

     ------------------------------------------------------------

8903 tgtccatggtaatacatccagattaaaatcgccagaaacaggaggagccaaaggagatca 8962
           ·      8920         ·      8940         ·      8960

     ------------------------------------------------------------

8963 accaatagagtccaccagttttgttgtagatagagagcaataatgagcaggatgaggtct 9022
           ·      8980         ·      9000         ·      9020

     ------------------------------------------------------------

9023 aggaagaaggctaggaagaaggccaaaagctgccagatggtggcaccaagtcgccagagc 9082
           ·      9040         ·      9060         ·      9080

 467 --------------------gatacc----------------------------------- 472
                          ||||||
9083 atctccaataagtagatccagatacctaagactgcgttgaaaaaagagtgttagggttgg 9142
           ·      9100         ·      9120         ·      9140
```

```
     ------------------------------------------------------------

9143 aaaagtgggggtgtggtaaataattcctagggaatgttagatcttaccaagtaagcaccc 9202
            ·       9160          ·       9180          ·       9200

     ------------------------------------------------------------

9203 gaagatgaacagcacaattccaaggaacaatgcctgtccgtgcaaattccagagagcgat 9262
            ·       9220          ·       9240          ·       9260

     ------------------------------------------------------------

9263 gagcaggagggtgactggggaaagaggagaaagtgcgttagagaaggaagagtaagggaa 9322
            ·       9280          ·       9300          ·       9320

     ------------------------------------------------------------

9323 aggggtgtggggcaaagggtgtaatacttactcatcagtaggagtatacaaagggctcc 9382
            ·       9340          ·       9360          ·       9380

     ------------------------------------------------------------

9383 aagtggacagagaaggtctcttctgaagataaagatgatcaaaattataattataagcat 9442
            ·       9400          ·       9420          ·       9440
                      480           ·       500           ·       520
 473 ----------gggt-caggtggcgggggacaaccccacgaca-ccgccccacgcggggca 520
               |  |  |||    || || |  | |||   ||  | | |    ||  | |
9443 gagagcaaaggaatagaggacaaggagggctcctccagtccagtcactcataacgatgta 9502
            ·       9460          ·       9480          ·       9500

 521 cgtaagaaacagtagccc---------------------------------------- 538
     |     ||||||||||xx
9503 cagccaaaacagtagcgccaagaggaggagaaggagagcaaggcctagggaagaggagag 9562
            ·       9520          ·       9540          ·       9560

     ------------------------------------------------------------

9563 gggggtcctcgaggggccgtcgcgggcccggtgggcccctctcaaggtcgtgttccat 9622
            ·       9580          ·       9600          ·       9620

     ------------------------------------------------------------

9623 cctcagggcagtgtgtcaggagcaaggcagttgaggaaagaaggggcagagcagtgtga 9682
            ·       9640          ·       9660          ·       9680

     ------------------------------------------------------------

9683 gaggcttatgtagggcggctacgtcagagtaacgcgtgtttcttgggatgtaggcccggg 9742
            ·       9700          ·       9720          ·       9740

     ------------------------------------------------------------

9743 gggatttgcggggtctgccggaggcagtacgggtacagatttcccgaaagcggcggtgtg 9802
            ·       9760          ·       9780          ·       9800

     ------------------------------------------------------------

9803 tgtgtgcatgtaagcgtagaaaggggaagtagaaagcgtgtgtttgtgttagaaaagcgg 9862
            ·       9820          ·       9840          ·       9860

     ------------------------------------------------------------

9863 gtccccgggggcaagctgtgggaatgcggtggccaagtgcaacaggaaatggaaaggca 9922
            ·       9880          ·       9900          ·       9920
```

```
9923 gtgcggcaatcagaaggggagtgcgtagtgttgtgggaagcggcagtgtaatctgcaca 9982
              9940          9960          9980

9983 aagaggcgcggggcgcgcaacgttgggaggtcgttggcggcaggcgggaggccgtgcttt 10042
             10000         10020         10040

10043 agggggggttcaggtgaggcaaggctgtggggtaaccgtaggggaggcgggtgaggcggct 10102
             10060         10080         10100

10103 aagagggctaagggtcggcgggtgacgaagcagcagacggcggatatgggaatttcagaa 10162
             10120         10140         10160

10163 tgaggtggcggattcaggcgaaaagggtgtgggctgtgcgagtgtcatgaggcaggcgcg 10222
             10180         10200         10220

10223 gaaagtcgctgcggcttgctggggcatgggggggccgcgcattcctggaaaaagtggaggg 10282
             10240         10260         10280

10283 ggcgtggccttcccccgcggccccccagccccccccgcacagagcggcgctacggcgggcg 10342
             10300         10320         10340

10343 ggcggcgggggggtcggggtccgcgggctccgggggctgcgggcggtggatggcggcggac 10402
             10360         10380         10400

10403 gttccggggatcgggggggtcgggggggcgccgcgcgggcgcagccatgcgtgaccgtgat 10462
             10420         10440         10460

10463 gaggggggcagggtcgcaggggggtgtgtctggtgggggcgggagcgggggggcggcgcggga 10522
             10480         10500         10520

10523 gcctgcacgccgttggagggtagaatgacaggggggcggggacagagaggcggtcgcgccc 10582
             10540         10560         10580

10583 ccggccgcgccagccaagcccccaagggggggcggggagcgggcaatggagcgtgacgaag 10642
             10600         10620         10640

10643 ggccccagggctgaccccggcaaacgtgacccggggctccggggtgacccagccaagcgt 10702
             10660         10680         10700
```

```
      ------------------------------------------------------------

10703 gaccaaggggcccgtgggtgacacaggcaaccctgacaaaggccccccaggaaagacccc 10762
            .      10720          .      10740          .      10760

      ------------------------------------------------------------

10763 cgggggggcatcgggggggtggggcatgggggggccgcgcattcctggaaaaagtggaggggg 10822
            .      10780          .      10800          .      10820

      ------------------------------------------------------------

10823 cgtggccttcccccgcggcccccccagcccccccgcacagagcggcgctacggcgggcgggg 10882
            .      10840          .      10860          .      10880

      ------------------------------------------------------------

10883 cggcgggggggtcggggtccgcgggctccgggggctgcgggcggtggatggcggcggacgt 10942
            .      10900          .      10920          .      10940

      ------------------------------------------------------------

10943 tccggggatcgggggggtcgggggggcgccgcgcgggcgcagccatgcgtgaccgtgatga 11002
            .      10960          .      10980          .      11000

      ------------------------------------------------------------

11003 gggggcagggtcgcaggggggtgtgtctggtgggggcgggagcgggggggcggcgcgggagc 11062
            .      11020          .      11040          .      11060

      ------------------------------------------------------------

11063 ctgcacgccgttggagggtagaatgacaggggggcggggacagagaggcggtcgcgccccc 11122
            .      11080          .      11100          .      11120

      ------------------------------------------------------------

11123 ggccgcgccagccaagcccccaagggggggcggggagcgggcaatggagcgtgacgaaggg 11182
            .      11140          .      11160          .      11180

      ------------------------------------------------------------

11183 ccccagggctgaccccggcaaacgtgacccggggctccggggtgacccagccaagcgtga 11242
            .      11200          .      11220          .      11240

      ------------------------------------------------------------

11243 ccaaggggcccgtgggtgacacaggcaaccctgacaaaggccccccaggaaagacccccg 11302
            .      11260          .      11280          .      11300

      ------------------------------------------------------------

11303 tggggcatgggggggccgcgcattcctggaaaaagtggaggggggcgtggccttcccccgcg 11362
            .      11320          .      11340          .      11360

      ------------------------------------------------------------

11363 gcccccccagcccccccgcacagagcggcgctacggcgggcgggcggcgggggggtcggggt 11422
            .      11380          .      11400          .      11420

      ------------------------------------------------------------

11423 ccgcgggctccgggggctgcgggcggtggatggcggcggacgttccggggatcgggggggg 11482
            .      11440          .      11460          .      11480
```

```
      ------------------------------------------------------------

11483 tcgggggcgccgcgcgggcgcagccatgcgtgaccgtgatgaggggcagggtcgcagg 11542
            .     11500         .     11520         .     11540

      ------------------------------------------------------------

11543 gggtgtgtctggtgggggcgggagcgggggcggcgcgggagcctgcacgccgttggagg 11602
            .     11560         .     11580         .     11600

      ------------------------------------------------------------

11603 gtagaatgacaggggcggggacagagaggcggtcgcgcccccggccgcgccagccaagc 11662
            .     11620         .     11640         .     11660

      ------------------------------------------------------------

11663 ccccaaggggggcggggagcgggcaatggagcgtgacgaagggccccagggctgaccccg 11722
            .     11680         .     11700         .     11720

      ------------------------------------------------------------

11723 gcaaacgtgacccggggctccggggtgacccagccaagcgtgaccaaggggcccgtgggt 11782
            .     11740         .     11760         .     11780

      ------------------------------------------------------------

11783 gacacaggcaaccctgacaaaggcccccccaggaaagacccccgggggggcatcgggggggtg 11842
            .     11800         .     11820         .     11840

      ------------------------------------------------------------

11843 gggcatgggggggccgcgcattcctggaaaaagtggaggggggcgtggccttcccccgcggc 11902
            .     11860         .     11880         .     11900

      ------------------------------------------------------------

11903 cccccagcccccccgcacagagcggcgctacggcgggcgggcggcgggggggtcggggtcc 11962
            .     11920         .     11940         .     11960

      ------------------------------------------------------------

11963 gcgggctccgggggctgcgggcggtggatggcggcggacgttccggggatcgggggggtc 12022
            .     11980         .     12000         .     12020

      ------------------------------------------------------------

12023 gggggggcgccgcgcgggcgcagccatgcgtgaccgtgatgaggggggcagggtcgcagggg 12082
            .     12040         .     12060         .     12080

      ------------------------------------------------------------

12083 gtgtgtctggtgggggcgggagcgggggggcggcgcgggagcctgcacgccgttggagggt 12142
            .     12100         .     12120         .     12140

      ------------------------------------------------------------

12143 agaatgacaggggcggggacagagaggcggtcgcgcccccggccgcgccagccaagccc 12202
            .     12160         .     12180         .     12200

      ------------------------------------------------------------

12203 ccaaggggggcggggagcgggcaatggagcgtgacgaagggccccagggctgaccccggc 12262
            .     12220         .     12240         .     12260
```

```
        ------------------------------------------------------------

12263 aaacgtgacccggggctccggggtgacccagccaagcgtgaccaaggggcccgtgggtga 12322
               .        12280        .        12300        .        12320

        ------------------------------------------------------------

12323 cacaggcaaccctgacaaaggccccccaggaaagaccccccgggggcatcggggggggtg 12382
               .        12340        .        12360        .        12380

        ------------------------------------------------------

12383 ttggcgggggcatgggggggtcggatttcgcccttattgccctgtttagaattc       12436
               .        12400        .        12420        .
```

% Identity = 2.3 (291/12474)

///

9310-13DVCTDV SEQ ID NO 1.xdna x Bankier et al. EcoRI Dhet fragment complement.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

DNA sequence 12436 bp gaattctaaaca ... ctttgagaattc linear

Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 8. Comparison of nucleotide sequence of SEQ ID NO:1 with the complement of the nucleotide sequence of Fig. 2 of Bankier et al.

```
   ------------------------------------------------------------

   1 gaattctaaacagggcaataagggcgaaatccgacccccccatgcccccgccaacacccc 60
               .         20         .         40         .         60

   ------------------------------------------------------------

  61 ccccgatgcccccgggggtctttcctgggggggcctttgtcagggttgcctgtgtcaccc 120
               .         80         .        100         .        120

   ------------------------------------------------------------

 121 acgggcccccttggtcacgcttggctgggtcaccccggagccccgggtcacgtttgccggg 180
               .        140         .        160         .        180

   ------------------------------------------------------------

 181 gtcagccctggggcccttcgtcacgctccattgcccgctccccgcccccccttgggggctt 240
               .        200         .        220         .        240

   ------------------------------------------------------------

 241 ggctggcgcggccgggggcgcgaccgcctctctgtccccgcccccctgtcattctaccctc 300
               .        260         .        280         .        300

   ------------------------------------------------------------

 301 caacggcgtgcaggctcccgcgccgcccccccgctcccgcccccaccagacacacccccctg 360
               .        320         .        340         .        360

   ------------------------------------------------------------

 361 cgaccctgcccccctcatcacggtcacgcatggctgcgcccgcgcggcgcccccccgacccc 420
               .        380         .        400         .        420

   ------------------------------------------------------------

 421 cccgatccccggaacgtccgccgccatccaccgcccgcagcccccggagcccgcggaccc 480
               .        440         .        460         .        480

   ------------------------------------------------------------

 481 cgacccccccgccgcccgcccgccgtagcgccgctctgtgcgggggggctgggggggccgcg 540
               .        500         .        520         .        540

   ------------------------------------------------------------

 541 ggggaaggccacgccccctccacttttttccaggaatgcgcggcccccccatgccccacccc 600
               .        560         .        580         .        600
```

```
    ------------------------------------------------------------

601 ccgatgccccccgggggtctttcctgggggggcctttgtcagggttgcctgtgtcacccac 660
          .       620         .       640         .       660

    ------------------------------------------------------------

661 gggccccttggtcacgcttggctgggtcaccccggagccccgggtcacgtttgccggggt 720
          .       680         .       700         .       720

    ------------------------------------------------------------

721 cagccctggggcccttcgtcacgctccattgcccgctccccgcccccctttgggggcttgg 780
          .       740         .       760         .       780

    ------------------------------------------------------------

781 ctggcgcggccgggggcgcgaccgcctctctgtccccgcccccctgtcattctaccctcca 840
          .       800         .       820         .       840

    ------------------------------------------------------------

841 acggcgtgcaggctcccgcgccgcccccgctcccgcccccaccagacacacccccctgcg 900
          .       860         .       880         .       900

    ------------------------------------------------------------

901 accctgcccccctcatcacggtcacgcatggctgcgcccgcgcggcgcccccgaccccc 960
          .       920         .       940         .       960

    ------------------------------------------------------------

961 cgatccccggaacgtccgccgccatccaccgcccgcagcccccggagcccgcggaccccg 1020
          .       980         .      1000         .      1020

    ------------------------------------------------------------

1021 acccccgccgcccgcccgccgtagcgccgctctgtgcgggggggctgggggccgcggg 1080
          .      1040         .      1060         .      1080

    ------------------------------------------------------------

1081 ggaaggccacgccccctccacttttccaggaatgcgcggcccccatgccccacggggg 1140
          .      1100         .      1120         .      1140

    ------------------------------------------------------------

1141 tctttcctgggggggcctttgtcagggttgcctgtgtcacccacgggccccttggtcacgc 1200
          .      1160         .      1180         .      1200

    ------------------------------------------------------------

1201 ttggctgggtcaccccggagccccgggtcacgtttgccggggtcagccctggggcccttc 1260
          .      1220         .      1240         .      1260

    ------------------------------------------------------------

1261 gtcacgctccattgcccgctccccgcccccctttgggggcttggctggcgcggccgggggc 1320
          .      1280         .      1300         .      1320

    ------------------------------------------------------------

1321 gcgaccgcctctctgtccccgcccccctgtcattctaccctccaacggcgtgcaggctccc 1380
          .      1340         .      1360         .      1380
```

```
     ------------------------------------------------------------

1381 gcgccgccccccgctcccgccccccaccagacacacccccctgcgaccctgccccctcatca 1440
               ·       1400          ·       1420          ·       1440

     ------------------------------------------------------------

1441 cggtcacgcatggctgcgcccgcgcggcgcccccccgaccccccccgatccccggaacgtcc 1500
               ·       1460          ·       1480          ·       1500

     ------------------------------------------------------------

1501 gccgccatccaccgcccgcagcccccggagcccgcggaccccgaccccccgccgcccgcc 1560
               ·       1520          ·       1540          ·       1560

     ------------------------------------------------------------

1561 cgccgtagcgccgctctgtgcgggggggctgggggccgcgggggaaggccacgccccct 1620
               ·       1580          ·       1600          ·       1620

     ------------------------------------------------------------

1621 ccacttttccaggaatgcgcggcccccccatgccccacccccgatgccccccgggggtc 1680
               ·       1640          ·       1660          ·       1680

     ------------------------------------------------------------

1681 tttcctgggggggcctttgtcagggttgcctgtgtcacccacgggcccccttggtcacgctt 1740
               ·       1700          ·       1720          ·       1740

     ------------------------------------------------------------

1741 ggctgggtcaccccggagccccgggtcacgtttgccggggtcagccctggggcccttcgt 1800
               ·       1760          ·       1780          ·       1800

     ------------------------------------------------------------

1801 cacgctccattgcccgctccccgcccccccttgggggcttggctggcgcggccgggggcgc 1860
               ·       1820          ·       1840          ·       1860

     ------------------------------------------------------------

1861 gaccgcctctctgtccccgcccccctgtcattctaccctccaacggcgtgcaggctcccgc 1920
               ·       1880          ·       1900          ·       1920

     ------------------------------------------------------------

1921 gccgcccccgctcccgcccccaccagacacacccccctgcgaccctgccccctcatcacg 1980
               ·       1940          ·       1960          ·       1980

     ------------------------------------------------------------

1981 gtcacgcatggctgcgcccgcgcggcgcccccccgaccccccccgatccccggaacgtccgc 2040
               ·       2000          ·       2020          ·       2040

     ------------------------------------------------------------

2041 cgccatccaccgcccgcagcccccggagcccgcggaccccgaccccccgccgcccgcccg 2100
               ·       2060          ·       2080          ·       2100

     ------------------------------------------------------------

2101 ccgtagcgccgctctgtgcgggggggctgggggccgcgggggaaggccacgcccccctcc 2160
               ·       2120          ·       2140          ·       2160
```

```
     ----------------------------------------------------------------

2161 acttttccaggaatgcgcggccccccatgccccagcaagccgcagcgactttccgcgcc 2220
               ·       2180          ·       2200          ·       2220

     ----------------------------------------------------------------

2221 tgcctcatgacactcgcacagcccacacccttttcgcctgaatccgccacctcattctga 2280
               ·       2240          ·       2260          ·       2280

     ----------------------------------------------------------------

2281 aattcccatatccgccgtctgctgcttcgtcacccgccgacccttagccctcttagccgc 2340
               ·       2300          ·       2320          ·       2340

     ----------------------------------------------------------------

2341 ctcacccgcctcccctacggttaccccacagccttgcctcacctgaaccccccctaaagca 2400
               ·       2360          ·       2380          ·       2400

     ----------------------------------------------------------------

2401 cggcctcccgcctgccgccaacgacctcccaacgttgcgcgccccgcgcctctttgtgca 2460
               ·       2420          ·       2440          ·       2460

     ----------------------------------------------------------------

2461 gattacactgccgcttcccacaacactacgcactccccttctgattgccgcactgcctt 2520
               ·       2480          ·       2500          ·       2520

     ----------------------------------------------------------------

2521 tccatttcctgttgcacttggccaccgcattcccacagcttgccccccggggacccgctt 2580
               ·       2540          ·       2560          ·       2580

     ----------------------------------------------------------------

2581 ttctaacacaaacacacgctttctacttccccttctacgcttacatgcacacacacacc 2640
               ·       2600          ·       2620          ·       2640

     ----------------------------------------------------------------

2641 gccgctttcgggaaatctgtacccgtactgcctccggcagaccccgcaaatccccccggg 2700
               ·       2660          ·       2680          ·       2700

     ----------------------------------------------------------------

2701 cctacatcccaagaaacacgcgttactctgacgtagccgccctacataagcctctcacac 2760
               ·       2720          ·       2740          ·       2760

     ----------------------------------------------------------------

2761 tgctctgccccctttctttcctcaactgccttgctcctgacacactgccctgaggatggaa 2820
               ·       2780          ·       2800          ·       2820

     ----------------------------------------------------------------

2821 cacgaccttgagaggggcccaccgggcccgcgacggcccccctcgaggacccccccctctcc 2880
               ·       2840          ·       2860          ·       2880

     ----------------------------------------------------------------

2881 tcttccctaggccttgctctccttctcctcctcttggcgctactgttttggctgtacatc 2940
               ·       2900          ·       2920          ·       2940
```

```
     ------------------------------------------------------------

2941 gttatgagtgactggactggaggagccctccttgtcctctattcctttgctctcatgctt 3000
               ·      2960          ·      2980          ·      3000

     ------------------------------------------------------------

3001 ataattataattttgatcatctttatcttcagaagagaccttctctgtccacttggagcc 3060
               ·      3020          ·      3040          ·      3060

     ------------------------------------------------------------

3061 ctttgtatactcctactgatgagtaagtattacacccttgccccacacccctttccct 3120
               ·      3080          ·      3100          ·      3120

     ------------------------------------------------------------

3121 tactcttccttctctaacgcactttctcctctttccccagtcaccctcctgctcatcgct 3180
               ·      3140          ·      3160          ·      3180

     ------------------------------------------------------------

3181 ctctggaatttgcacggacaggcattgttccttggaattgtgctgttcatcttcgggtgc 3240
               ·      3200          ·      3220          ·      3240

     ------------------------------------------------------------

3241 ttacttggtaagatctaacattccctaggaattatttaccacacccccactttccaacc 3300
               ·      3260          ·      3280          ·      3300

     ------------------------------------------------------------

3301 ctaacactctttttcaacgcagtcttaggtatctggatctacttattggagatgctctg 3360
               ·      3320          ·      3340          ·      3360

     ------------------------------------------------------------

3361 gcgacttggtgccaccatctggcagcttttggccttcttcctagccttcttcctagacct 3420
               ·      3380          ·      3400          ·      3420

     ------------------------------------------------------------

3421 catcctgctcattattgctctctatctacaacaaaactggtggactctattggttgatct 3480
               ·      3440          ·      3460          ·      3480

     ------------------------------------------------------------

3481 cctttggctcctcctgtttctggcgattttaatctggatgtattaccatggacaacgaca 3540
               ·      3500          ·      3520          ·      3540

     ------------------------------------------------------------

3541 cagtgatgaacaccaccacgatgactccctcccgcaccctcaacaagctaccgatgattc 3600
               ·      3560          ·      3580          ·      3600

     ------------------------------------------------------------

3601 tggccatgaatctgactctaactccaacgagggcagacaccacctgctcgtgagtggagc 3660
               ·      3620          ·      3640          ·      3660

     ------------------------------------------------------------

3661 cggcgacggacccccactctgctctcaaaacctaggcgcacctggaggtggtcctgacaa 3720
               ·      3680          ·      3700          ·      3720
```

```
     ------------------------------------------------------------

3721 tggcccacaggaccctgacaacactgatgacaatggcccacaggaccctgacaacactga 3780
               ·       3740          ·       3760          ·       3780

     ------------------------------------------------------------

3781 tgacaatggcccacatgacccgctgcctcaggaccctgacaacactgatgacaatggccc 3840
               ·       3800          ·       3820          ·       3840

     ------------------------------------------------------------

3841 acaggaccctgacaacactgatgacaatggcccacatgacccgctgcctcatagccctag 3900
               ·       3860          ·       3880          ·       3900

   1 -------------catgatggcacgccggct------------------------------ 18
                 x|||||||     ||  |
3901 cgactctgctggaaatgatggaggccctccacaattgacggaagaggttgaaaacaaagg 3960
               ·       3920          ·       3940          ·       3960

     ------------------------------------------------------------

3961 aggtgaccagggcccgcctttgatgacagacggaggcggcggtcatagtcatgattccgg 4020
               ·       3980          ·       4000          ·       4020

     ------------------------------------------------------------

4021 ccatggcggcggtgatccacaccttcctacgctgcttttgggttcttctggttccggtgg 4080
               ·       4040          ·       4060          ·       4080

     ------------------------------------------------------------

4081 agatgatgacgacccccacggcccagttcagctaagctactatgactaacctttctttac 4140
               ·       4100          ·       4120          ·       4140

     ------------------------------------------------------------

4141 ttctaggcattaccatgtcataggcttgcctgactgactctccctccatttactgggaat 4200
               ·       4160          ·       4180          ·       4200

     ------------------------------------------------------------

4201 gccttagctaatcaccttaactggcacacactcccttagccacactgtctgtctaggctg 4260
               ·       4220          ·       4240          ·       4260

     ------------------------------------------------------------

4261 aaaagccacattcatattctatttcaaaacaaggggaaaggaggacatgcgagaattggc 4320
               ·       4280          ·       4300          ·       4320

     ------------------------------------------------------------

4321 agacacctttacccagcccttaacacaccacacaggtagcaaggacccgggcgttgccag 4380
               ·       4340          ·       4360          ·       4380

     ------------------------------------------------------------

4381 actccgccaccaacgcccctgcgttgaacccacccctcctacacacatcagacctctgca 4440
               ·       4400          ·       4420          ·       4440

     ------------------------------------------------------------

4441 caacacaactaccaggcagatgaggccccttacttccacagggtactggcataccagcgg 4500
               ·       4460          ·       4480          ·       4500
```

```
      ------------------------------------------------------------

4501 gggaccacatacatccctgtctcccacccagtaactccagcaactttgctttccatcttg 4560
           .       4520          .       4540          .       4560

                        20
  19 ----------------------gcccaagcc---------------------------- 27
                           |||||||||
4561 tgccaatacacatttggattcagcccaagccacacctaactcatgccagcagaggcagga 4620
           .       4580          .       4600          .       4620

      ------------------------------------------------------------

4621 acacctgttgttgacacattctttgcgcataagcactttaatccctctctcacacccaga 4680
           .       4640          .       4660          .       4680

      ------------------------------------------------------------

4681 aactaagagctagcccaaaacctccacacctgtcctcgctcatctttccacattcctctg 4740
           .       4700          .       4720          .       4740

      ------------------------------------------------------------

4741 gccttctttccttgtccttactgtataaaagtccacgaaaacagctgtgcctcactctcg 4800
           .       4760          .       4780          .       4800

      ------------------------------------------------------------

4801 agatggtacacgtcctggagcgtgctttgctagagcagcagtcctctgcctgcggcctgc 4860
           .       4820          .       4840          .       4860

      ------------------------------------------------------------

4861 ccggctcttctacggagaccaggcctagccacccctgccccgaggacccagacgtcagca 4920
           .       4880          .       4900          .       4920

      ------------------------------------------------------------

4921 gactaagactactcctggtggtactctgtgtcctgtttggacttttatgcctgctcctca 4980
           .       4940          .       4960          .       4980

      ------------------------------------------------------------

4981 tctaagaagccaccatgcgaccgggtagaccactggctggattctacgctactctccgcc 5040
           .       5000          .       5020          .       5040

      ------------------------------------------------------------

5041 gttccttcagaagaatgtccaaaaggtcaaagaacaaggccaagaaggagcgtgtccccg 5100
           .       5060          .       5080          .       5100

      ------------------------------------------------------------

5101 tggaggaccgcccaccgactccgatgcccaccagccagcgactgatccgcagaaacgcgt 5160
           .       5120          .       5140          .       5160

      ------------------------------------------------------------

5161 tgggaggaggcgtccgccccgatgcggaggactgcatccaacgcttccacccccctggagc 5220
           .       5180          .       5200          .       5220

      ------------------------------------------------------------

5221 cagcgctgggggtgtcaacaaagaactttgacctgttgtccctgagatgtgaattgggat 5280
           .       5240          .       5260          .       5280
```

```
     ------------------------------------------------------------

5281 ggtgtggataacatctcccgctagatggcgcccttattattgatgtgacttgtgatgcaa 5340
               .       5300          .       5320          .       5340

     ------------------------------------------------------------

5341 taaataaaagtacagatagatggcactcttaccttcctctgcccgcttcttcgtatatgt 5400
               .       5360          .       5380          .       5400

     ------------------------------------------------------------

5401 gttgagatgagtcatcccgtggagagtagggaggggggagggagcccgtcattcccgtcgt 5460
               .       5420          .       5440          .       5460

     ------------------------------------------------------------

5461 gttgcaatcccaagtacagactttgatcttgggttcctagtggttgatagtccgagtgac 5520
               .       5480          .       5500          .       5520

     ------------------------------------------------------------

5521 ggtcgccattgccccaatatgggtcctcataaggcggtgggggctcttcattagattcac 5580
               .       5540          .       5560          .       5580

     ------------------------------------------------------------

5581 gttcctcatcgttcggtggggtgggggtgttcccagaagagccagaagcagatggatatt 5640
               .       5600          .       5620          .       5640

     ------------------------------------------------------------

5641 gggagttgtttccgccatcgtacccatccggatccccgccggggctaggggacccgcgc 5700
               .       5660          .       5680          .       5700

     ------------------------------------------------------------

5701 ccattggcaccatttctagggaccccatagctgcagcagcgactgcaaaaccggcctgaa 5760
               .       5720          .       5740          .       5760

     ------------------------------------------------------------

5761 atgctctctgagaaacaaggcgagagggatttacgccccagcaagcttcaggcaaaataa 5820
               .       5780          .       5800          .       5820

     ------------------------------------------------------------

5821 cctgtgatgctgaaaccgcggcgctacaggatcaccccaattggcaagacctgggcggag 5880
               .       5840          .       5860          .       5880

     ------------------------------------------------------------

5881 ttcctgagagcccaggggtctcgtgcaggtgtccccggggaattgttctccctgatcacc 5940
               .       5900          .       5920          .       5940

  28 ---caccc------------------------------------------------------ 32
        |||||
5941 gcccacccccgttttctccaaaccagcagctgtgacaatcttcacacactgctgctgtca 6000
               .       5960          .       5980          .       6000

     ------------------------------------------------------------

6001 cctggaactattttcccacggtgcccttccgcccattttcccacgagtcgcgaggctatc 6060
               .       6020          .       6040          .       6060
```

```
      ------------------------------------------------------------

6061  cacccgcaatgccacccccccaatgccacactaaaacaaggtgaaataggcaagtgcgtt 6120
                ·       6080          ·       6100          ·       6120

      ------------------------------------------------------------

6121  tattgcgacaagtatccagaaacataaaccccgtgggcttcctccttgtcatttttccca 6180
                ·       6140          ·       6160          ·       6180

      ------------------------------------------------------------

6181  acgcaggtcactggcaggtgccagggcttgggaagtgacaggtcaacagcaacagagagg 6240
                ·       6200          ·       6220          ·       6240

      ------------------------------------------------------------

6241  ctcccatccttttccttcataacaccgccatttgccgcagttggtgcgggctccacgccc 6300
                ·       6260          ·       6280          ·       6300

      ------------------------------------------------------------

6301  tcgggcatgagccactggacgtggggatggggaaatgcattcacggtgcatgtcacagta 6360
                ·       6320          ·       6340          ·       6360

      ------------------------------------------------------------

6361  aggacagagaagtctgggaactgagacctttcggagtggacagacagcgttagaggcttc 6420
                ·       6380          ·       6400          ·       6420

      ------------------------------------------------------------

6421  accacgctcaggtgttcctgcttggtgacctcggtctcgcccagtttcatgcggcacagg 6480
                ·       6440          ·       6460          ·       6480

      ------------------------------------------------------------

6481  tagttgccgtcatgggagatgttggcagcggtgactactaaaaagaaggtgttggcactt 6540
                ·       6500          ·       6520          ·       6540

      ------------------------------------------------------------

6541  ctgtggatatcaaagaagcccctgaaaggccactctataaagatgacatcgtggtgcatg 6600
                ·       6560          ·       6580          ·       6600

      ------------------------------------------------------------

6601  cgcccaataagcacctgctcctctcctgggcccagtttaaaccagctgacctcaatctct 6660
                ·       6620          ·       6640          ·       6660

        33  --------------------tccag------------------------------------ 37
                                 |||||
6661  ggaccgaggctcaccctcctccagtaggaggtcagggtgactcgctcacccaagaaagcg 6720
                ·       6680          ·       6700          ·       6720

      ------------------------------------------------------------

6721  gtgacagcctggccggcggccacacaggaggccaacaggaggagctgagcgatgaacctg 6780
                ·       6740          ·       6760          ·       6780

      ------------------------------------------------------------

6781  gccattgctctggactctcctcacccaggcctcgcgtcttatactattctgccacgccca 6840
                ·       6800          ·       6820          ·       6840
```

```
      ------------------------------------------------------------

6841  ttttatcatataagcctgaagcccgtgagctggcctgacgagaccatgaggccagccaag 6900
                .       6860          .       6880          .       6900

      ------------------------------------------------------------

6901  tctacagattctgtgtttgtgaggaccccggtcgaggcgtgggtcgcgccctcgccgccg 6960
                .       6920          .       6940          .       6960

      ------------------------------------------------------------

6961  gacgacaaggtggctgagtccagctacctcatgttcagggccatgtacgcggtgttcacc 7020
                .       6980          .       7000          .       7020

      ------------------------------------------------------------

7021  cgggatgagaaagacctgcctttgccagccctggtcctctgccggctcatcaaggcctcc 7080
                .       7040          .       7060          .       7080

      ------------------------------------------------------------

7081  ctgaggaaggataggaagctgtacgcggagctggcctgcaggacagccgacatcgggggc 7140
                .       7100          .       7120          .       7140

      ------------------------------------------------------------

7141  aaagacacgcacgtacggctcatcatcagcgtcctgcgcgcagtgtacaacgaccactac 7200
                .       7160          .       7180          .       7200

      ------------------------------------------------------------

7201  gactactggtcgcggctcagggtggtgctgtgctacacagtggtgtttgcggtgcgaaac 7260
                .       7220          .       7240          .       7260

      ------------------------------------------------------------

7261  tacctggatgaccacaagagcgccgccttcgtgctgggggcaatcgcccactacctggcc 7320
                .       7280          .       7300          .       7320
                              40
  38  ----------------------gggaggctggaggcggattt-------------------- 57
                             | ||||||||xxxxxxxxxx
7321  ctctatcgcagactctggtttgcgaggctgggcggcatgccaagatcgctgagacgtcag 7380
                .       7340          .       7360          .       7380

      ------------------------------------------------------------

7381  ttccccgtgacgtgggccctggccagcctgactgacttcctgaaatctttgtaaatgaat 7440
                .       7400          .       7420          .       7440

      ------------------------------------------------------------

7441  aaacagtgggtgttgcgtgatgagtaaagtgtaacatttaatgtgggactgggaggccgg 7500
                .       7460          .       7480          .       7500

      ------------------------------------------------------------

7501  ggcgataccttgggcatcatgcagggtgcacagactagcgaggataatctgggcagccag 7560
                .       7520          .       7540          .       7560

      ------------------------------------------------------------

7561  agccagccgggtccgtgcggctacatctacttttacccccctggccacctaccctcttagg 7620
                .       7580          .       7600          .       7620
```

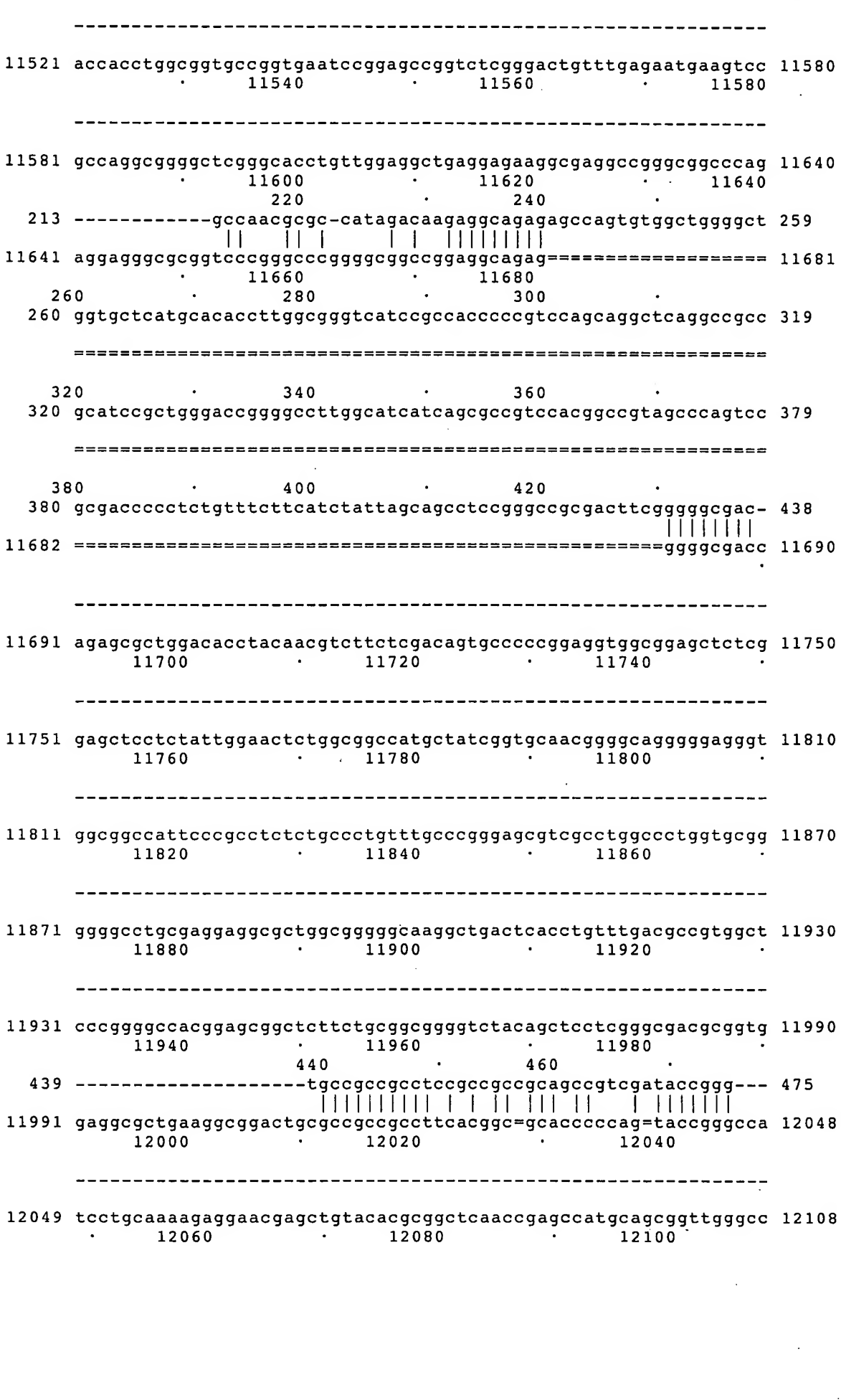
```
     ------------------------------------------------------------

7621 gaggtggccacactggggaccggctacgcgggccacaggtgcctgacggtgccgctcctt 7680
                 ·     7640         ·     7660         ·     7680

     ------------------------------------------------------------

7681 tgcggcatcaccgtggagccgggcttcagcatcaatgtcaaggctctgcacaggaggccc 7740
                 ·     7700         ·     7720         ·     7740

     ------------------------------------------------------------

7741 gaccccaactgcgggctcctacgcgctacctcctatcacagggacatctacgtgttccac 7800
                 ·     7760         ·     7780         ·     7800

     ------------------------------------------------------------

7801 aatgcccatatggttcccccatctttgaggggccgggtctcgaggccctctgtggcgag 7860
                 ·     7820         ·     7840         ·     7860

     ------------------------------------------------------------

7861 accagggaggtgtttgggtacgacgcctacagcgccctaccgagggaaagctccaagccg 7920
                 ·     7880         ·     7900         ·     7920

     ------------------------------------------------------------

7921 ggggacttcttccccgaagggctagatccctctgcctacctgggggcggtggcaataacc 7980
                 ·     7940         ·     7960         ·     7980

     ------------------------------------------------------------

7981 gaggccttcaaggagcgactctacagcggaaacctggtggccattccatcgttaaaacag 8040
                 ·     8000         ·     8020         ·     8040

     ------------------------------------------------------------

8041 gaggtagcggtggggcagtctgcgagcgttagggtcccgctctacgacaaggaggtgttc 8100
                 ·     8060         ·     8080         ·     8100

     ------------------------------------------------------------

8101 ccagagggcgtgccccagctccgccagttttacaactcggacctcagccgctgcatgcac 8160
                 ·     8120         ·     8140         ·     8160

     ------------------------------------------------------------

8161 gaggcgctgtacaccgggctggcgcaggcgctgcgcgtccgacgggtgggcaagctggtg 8220
                 ·     8180         ·     8200         ·     8220

     ------------------------------------------------------------

8221 gagctgctggagaagcagagcctgcaggaccaggccaaggtggccaaggtggccccctc 8280
                 ·     8240         ·     8260         ·     8280

     ------------------------------------------------------------

8281 aaggagttcccagcctcaaccatcagtcacccggactcgggagccttaatgattgtggac 8340
                 ·     8300         ·     8320         ·     8340

     ------------------------------------------------------------

8341 agcgcggcatgcgagctggcggtgagctacgcacccgccatgctggaggcctcgcacgag 8400
                 ·     8360         ·     8380         ·     8400
```

```
                                                                60
  58 ------------------------------------------------------tccaga-  63
                                                       ||||||
8401 accccggccagcctcaactacgactcgtggcccctgtttgccgactgtgagggtccagag 8460
              .       8420          .       8440          .       8460

     ------------------------------------------------------------

8461 gcccgtgtggctgcgttacaccgatataatgccagcctggccccccacgtgtccacgcag 8520
              .       8480          .       8500          .       8520

     ------------------------------------------------------------

8521 atctttgccaccaattccgtcctctacgtctcggggtctcgaagtcaaccggtcagggc 8580
              .       8540          .       8560          .       8580

  64 -----------------cagt---------------------------------------  67
                      ||||
8581 aaggagagtctctttaacagtttctacatgacccacggcctggggaccctgcaggaggg 8640
              .       8600          .       8620          .       8640

  68 --------------------cccctgcttc------------------------------  77
                          |||||||||
8641 acctgggacccctgccgccgaccctgcttctcgggctggggtggccagacgtgaccgga 8700
              .       8660          .       8680          .       8700

     ------------------------------------------------------------

8701 accaacggtccgggaaactacgctgtggagcacctggtctatgcggcctccttctcgccc 8760
              .       8720          .       8740          .       8760

     ------------------------------------------------------------

8761 aaccttcttgcccgctatgcctactacctgcagttttgccagggacagaagagctctctg 8820
              .       8780          .       8800          .       8820

     ------------------------------------------------------------

8821 accccggtgccggagacgggcagctacgtggcgggggcggccgccagtcccatgtgctcg 8880
              .       8840          .       8860          .       8880

     ------------------------------------------------------------

8881 ctctgcgagggccgggccccggccgtgtgcctgaacacgctcttctttaggctgagggac 8940
              .       8900          .       8920          .       8940

     ------------------------------------------------------------

8941 cgcttcccccccgtcatgtccacgcagcggagggacccctatgtgatctcgggggcctcg 9000
              .       8960          .       8980          .       9000

     ------------------------------------------------------------

9001 ggctcctacaacgagacggactttttgggcaactttctcaacttcatcgataaggaggac 9060
              .       9020          .       9040          .       9060
                                   80          .         100
  78 ----------------------------cta-aatttcaag-agctgaaccagaa----  102
                                 ||| |   | |  | |||||||||||||
9061 gacgggcagcggccggacgacgagccccgctacacctactggcagctgaaccagaacctg 9120
              .       9080          .       9100          .       9120

     ------------------------------------------------------------

9121 ctggagcggctgtctcggctgggcatagacgctgaaggaaagctagagaaggagccccat 9180
              .       9140          .       9160          .       9180
```

```
     ------------------------------------------------------------
9181 ggcccgcgtgactttgtcaagatgttcaaggacgtggatgcggcggtggacgccgaagtg 9240
              .        9200       .       9220        .      9240

     ------------------------------------------------------------
9241 gtccagtttatgaacagcatggccaagaacaacatcacctacaaggacctggtcaagagc 9300
              .        9260       .       9280        .      9300

     ------------------------------------------------------------
9301 tgctaccacgtgatgcagtactcgtgcaacccctttgcgcagcccgcctgccccatcttc 9360
              .        9320       .       9340        .      9360
                                                   .      120
 103 ----------------------------------------taatctccccaatgatgttt 122
                                               |||||||      || | |
9361 acccagctgttttaccgctcactgctgaccatcctgcaggacatctccctgcccatctgt 9420
              .        9380       .       9400        .      9420

 123 ttcgg------------------------------------------------------- 127
      |  |
9421 atgtgctatgagaatgacaaccccgggcttggccagagcccccccagagtggctaaaggt 9480
              .        9440       .       9460        .      9480

     ------------------------------------------------------------
9481 cactaccagacgctgtgcaccaactttaggagcctggccatcgacaaggggggtcctcacg 9540
              .        9500       .       9520        .      9540

     ------------------------------------------------------------
9541 gccaaggaggccaaggtggtgcatggggagcccacctgcgacctgccagacctggacgcg 9600
              .        9560       .       9580        .      9600

     ------------------------------------------------------------
9601 gccctgcagggccgggtgtacggccggcggctgcctgtgcgcatgtccaaggtgctgatg 9660
              .        9620       .       9640        .      9660

     ------------------------------------------------------------
9661 ctgtgccccaggaacatcaagatcaagaacagggtggtcttcacgggggagaatgccgcc 9720
              .        9680       .       9700        .      9720

     ------------------------------------------------------------
9721 ctccagaacagcttcatcaagtccactaccaggagggagaactacatcatcaacgggccc 9780
              .        9740       .       9760        .      9780

     ------------------------------------------------------------
9781 tacatgaaattcctcaacacctaccacaagaccctattcccggacactaagctctcaagc 9840
              .        9800       .       9820        .      9840

     ------------------------------------------------------------
9841 ctgtacctgtggcacaacttttccaggcggcgctcggtccctgtccccagcggggccagc 9900
              .        9860       .       9880        .      9900

     ------------------------------------------------------------
9901 gcggaggagtactctgacctggccctctttgtggacggggctcccgggcccacgaagag 9960
              .        9920       .       9940        .      9960
```

```
  128 ----------------------------------------------------gaggctcaaa 137
                                                         |||||||||
 9961 agcaacgtcatagatgtggtgcctggcaacctggtcacttacgccaagcagaggctcaac 10020
            .      9980          .     10000          .     10020
      140           .       160          .              180
  138 gaagtta-cctg--gtatttctgacat-cccagttctgc-t-a-c--gaag-agt-ac-g 185
       | |   | ||||  |    | | | ||  |||||||||| | | | |  ||   ||  || |
10021 aacgccatcctgaaggcgtgcggccagacccagttctacatcagcctgattcagggactg 10080
            .     10040          .     10060          .     10080
          .       200
  186 -tgcagaggacttttggggt---------------------------------------- 204
       ||| ||||||      |||
10081 gtgccgaggacgcagtcggtgcccgcccgtgactacccccacgtactgggcacgcgggcg 10140
            .     10100          .     10120          .     10140

      ------------------------------------------------------------

10141 gtggagtcggcagcggcctacgcggaggccacctcctcccttactgcgaccacggtggtc 10200
            .     10160          .     10180          .     10200

      ------------------------------------------------------------

10201 tgcgcggccacagactgtcttagccaggtctgcaaggcccgtccggttgtcacgctgcca 10260
            .     10220          .     10240          .     10260

      ------------------------------------------------------------

10261 gtgaccatcaacaagtacacgggggtcaacggcaacaaccagatattccaggccgggaac 10320
            .     10280          .     10300          .     10320

      ------------------------------------------------------------

10321 ctgggatactttatgggccggggcgtggacaggaacctgctgcaggcccccggggctggg 10380
            .     10340          .     10360          .     10380

      ------------------------------------------------------------

10381 ctgcgcaagcaggccgggggctcttccatgcggaagaagtttgtctttgccacccccacc 10440
            .     10400          .     10420          .     10440

      ------------------------------------------------------------

10441 ctagggttgaccgtgaagcgccggacccaagccgcgaccacatatgagattgagaacatc 10500
            .     10460          .     10480          .     10500

      ------------------------------------------------------------

10501 agggctggcctggaggccattatatcacaaaaacaggaggaagactgtgtgtttgatgtg 10560
            .     10520          .     10540          .     10560

      ------------------------------------------------------------

10561 gtgtgcaaccttgtggatgccatgggcgaggcatgcgcctcgctgactagggacgacgcg 10620
            .     10580          .     10600          .     10620

      ------------------------------------------------------------

10621 gagtacttattgggccgcttctccgtcctggcggacagcgtcctagaaaccctggcgacc 10680
            .     10640          .     10660          .     10680

      ------------------------------------------------------------

10681 attgcctccagcgggatagagtggacggcggaggccgctcgggactttctggagggagtg 10740
            .     10700          .     10720          .     10740
```

```
      ------------------------------------------------------------

10741 tggggtgggcccggggcagcccaggacaactttatcagcgtggccgagccggtcagcacc 10800
            .     10760         .     10780         .     10800

  205 ---------gcctcggc------------------------------------------- 212
              ||||||||
10801 gcgtcgcaggcctcggccgggctgctgctgggtggaggagggcagggctccgggggcaga 10860
            .     10820         .     10840         .     10860

      ------------------------------------------------------------

10861 cgcaagcgccgtctggccaccgttctccccggactcgaggtctagagacccctggggcgg 10920
            .     10880         .     10900         .     10920

      ------------------------------------------------------------

10921 cgatgtcggggctgctggcggcggcgtacagccaggtgtacgccctggcggttgagctga 10980
            .     10940         .     10960         .     10980

      ------------------------------------------------------------

10981 gcgtgtgcacccggctggacccccggagtctggacgtggctgcggtggtgcgcaacgccg 11040
            .     11000         .     11020         .     11040

      ------------------------------------------------------------

11041 gcctgctggccgagctggaggccatcctccttccccgtttgagacggcagaatgaccgtg 11100
            .     11060         .     11080         .     11100

      ------------------------------------------------------------

11101 catgcagcgccctgtccctggagctggtgcacctgctagagaactcgagagaggcctctg 11160
            .     11120         .     11140         .     11160

      ------------------------------------------------------------

11161 ccgcgctgctcgcccctggtagaaagggtacccgggtcccgcctctccgtaccccctcag 11220
            .     11180         .     11200         .     11220

      ------------------------------------------------------------

11221 tcgcgtactctgtggagttttacgggggcataaagtcgatgtaagtttgtgcctaataa 11280
            .     11240         .     11260         .     11280

      ------------------------------------------------------------

11281 atgacatagagattttaatgaagagaatcaatagcgtgttttattgcatgtctcacacca 11340
            .     11300         .     11320         .     11340

      ------------------------------------------------------------

11341 tggggctggagagcctggaacgggccctggatctgctgggccgctttcggggcgtaagtc 11400
            .     11360         .     11380         .     11400

      ------------------------------------------------------------

11401 ccatcccagacccgcgcctctacatcacctctgtgccctgctggcgctgtgtgggcgagc 11460
            .     11420         .     11440         .     11460

      ------------------------------------------------------------

11461 tgatggttctgcccaaccacggcaacccttccacggcagaggggacccacgtctcctgta 11520
            .     11480         .     11500         .     11520
```


11521 accacctggcggtgccggtgaatccggagccggtctcgggactgtttgagaatgaagtcc 11580
11581 gccaggcggggctcgggcacctgttggaggctgaggagaaggcgaggccgggcggcccag 11640
213 ----------gccaacgcgc-catagacaagaggcagagagccagtgtggctggggct 259
11641 aggagggcgcggtcccgggcccggggcggccggaggcagag========== 11681
260 ggtgctcatgcacaccttggcgggtcatccgccacccccgtccagcaggctcaggccgcc 319
320 gcatccgctgggaccggggccttggcatcatcagcgccgtccacggccgtagcccagtcc 379
380 gcgacccccctctgtttcttcatctattagcagcctccgggccgcgacttcgggggcgac- 438
11682 ==========ggggcgacc 11690
11691 agagcgctggacacctacaacgtcttctcgacagtgcccccggaggtggcggagctctcg 11750
11751 gagctcctctattggaactctggcggccatgctatcggtgcaacggggcaggggagggt 11810
11811 ggcggccattcccgcctctctgccctgtttgcccgggagcgtcgcctggccctggtgcgg 11870
11871 ggggcctgcgaggaggcgctggcgggggcaaggctgactcacctgtttgacgccgtggct 11930
11931 cccggggccacggagcggctcttctgcggcggggtctacagctcctcgggcgacgcggtg 11990
439 ----------tgccgccgcctccgccgccgcagccgtcgataccggg--- 475
11991 gaggcgctgaaggcggactgcgccgccgccttcacggc=gcacccccag=taccgggcca 12048
12049 tcctgcaaaagaggaacgagctgtacacgcggctcaaccgagccatgcagcggttgggcc 12108

```
         ------------------------------------------------------------

   12109 gaggcgaggaggaggcgtcccgggagagcccggaggtgccccggccggctggggcacgag 12168
            .      12120       .      12140       .      12160

     476 ----------------------------------------------------------tc 477
                                                                   |
   12169 agcccggcccgtccggcgccctctcggacgcgctcaagcgcaaggagcagtacctgcgcc 12228
            .      12180       .      12200       .      12220
         480
     478 aggtggc----------------------------------------------------- 484
         |||||||
   12229 aggtggccaccgagggtctggccaagctgcagtcctgcctggcgcaacagagcgagaccc 12288
            .      12240       .      12260       .      12280
                                                .         500         .
     485 ----------------------------gggggacaacccc-ac-gacacc-g-ccc-ca 511
                                     |||||||     | || |   | | | ||| ||
   12289 tgaccgagaccctgtgcctgcgcgtctgggggggacgtggtctactgggagctggcccgca 12348
            .      12300       .      12320       .      12340
                   520         .
     512 cgcggggc-acgtaagaaacagtagccc-------------------------------- 538
          |||   | || |     ||||  |  |
   12349 tgcgcaaccacttcctctacagacgggccttcgtctcgggtccctgggaggacaggcgcg 12408
            .      12360       .      12380       .      12400

         ----------------------------

   12409 ccggcgagggtgccgcctttgagaattc                                 12436
            .      12420       .

  % Identity =   1.9 (243/12628)

///
```

DNA Strider 1.4f6 ### Friday, July 13, 2007 5:22:43 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 3.xdna x Bankier et al. EcoRI Dhet fragment.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

DNA sequence 12436 bp gaattctcaaag ... tgtttagaattc linear

Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 9. Comparison of nucleotide sequence of SEQ ID NO:3 with the nucleotide sequence of Fig. 2 of Bankier et al.

```
    ------------------------------------------------------------

  1 gaattctcaaaggcggcaccctcgccggcgcgcctgtcctcccagggacccgagacgaag 60
              .         20          .         40          .         60

    ------------------------------------------------------------

 61 gcccgtctgtagaggaagtggttgcgcatgcgggccagctcccagtagaccacgtccccc 120
              .         80          .        100          .        120

    ------------------------------------------------------------

121 cagacgcgcaggcacagggtctcggtcagggtctcgctctgttgcgccaggcaggactgc 180
              .        140          .        160          .        180

    ------------------------------------------------------------

181 agcttggccagaccctcggtggccacctggcgcaggtactgctccttgcgcttgagcgcg 240
              .        200          .        220          .        240

    ------------------------------------------------------------

241 tccgagagggcgccggacgggccgggctctcgtgccccagccggccggggcacctccggg 300
              .        260          .        280          .        300

    ------------------------------------------------------------

301 ctctcccgggacgcctcctcctcgcctcggcccaaccgctgcatggctcggttgagccgc 360
              .        320          .        340          .        360

    ------------------------------------------------------------

361 gtgtacagctcgttcctcttttgcaggatggcccggtactgggggtgcgccgtgaaggcg 420
              .        380          .        400          .        420

    ------------------------------------------------------------

421 gcggcgcagtccgccttcagcgcctccaccgcgtcgcccgaggagctgtagaccccgccg 480
              .        440          .        460          .        480

    ------------------------------------------------------------

481 cagaagagccgctccgtggccccgggagccacggcgtcaaacaggtgagtcagccttgcc 540
              .        500          .        520          .        540

    ------------------------------------------------------------

541 cccgccagcgcctcctcgcaggcccccgcaccagggccaggcgacgctcccgggcaaac 600
              .        560          .        580          .        600
```

```
    ------------------------------------------------------------

601 agggcagagaggcgggaatggccgccaccctccccctgccccgttgcaccgatagcatgg 660
          .       620         .       640         .       660

    ------------------------------------------------------------

661 ccgccagagttccaatagaggagctccgagagctccgccacctccgggggcactgtcgag 720
          .       680         .       700         .       720

    ------------------------------------------------------------

721 aagacgttgtaggtgtccagcgctctggtcgcccccctctgcctccggccgccccgggccc 780
          .       740         .       760         .       780

    ------------------------------------------------------------

781 gggaccgcgccctcctctgggccgcccggcctcgccttctcctcagcctccaacaggtgc 840
          .       800         .       820         .       840

    ------------------------------------------------------------

841 ccgagccccgcctggcggacttcattctcaaacagtcccgagaccggctccggattcacc 900
          .       860         .       880         .       900

    ------------------------------------------------------------

901 ggcaccgccaggtggttacaggagacgtgggtcccctctgccgtggaagggttgccgtgg 960
          .       920         .       940         .       960

    ------------------------------------------------------------

961 ttgggcagaaccatcagctcgcccacacagcgccagcagggcacagaggtgatgtagagg 1020
          .       980         .      1000         .      1020

    ------------------------------------------------------------

1021 cgcgggtctgggatgggacttacgccccgaaagcggcccagcagatccagggcccgttcc 1080
          .      1040         .      1060         .      1080

    ------------------------------------------------------------

1081 aggctctccagccccatggtgtgagacatgcaataaaacacgctattgattctcttcatt 1140
          .      1100         .      1120         .      1140

    ------------------------------------------------------------

1141 aaaatctctatgtcatttattaggcacaaacttacatcgactttatgccccccgtaaaac 1200
          .      1160         .      1180         .      1200

    ------------------------------------------------------------

1201 tccacagagtacgcgactgaggggtacggagaggcgggacccgggtacccttttctacca 1260
          .      1220         .      1240         .      1260

    ------------------------------------------------------------

1261 ggggcgagcagcgcggcagaggcctctctcgagttctctagcaggtgcaccagctccagg 1320
          .      1280         .      1300         .      1320

    ------------------------------------------------------------

1321 gacagggcgctgcatgcacggtcattctgccgtctcaaacggggaaggaggatggcctcc 1380
          .      1340         .      1360         .      1380
```

```
     ------------------------------------------------------------

1381 agctcggccagcaggccggcgttgcgcaccaccgcagccacgtccagactccgggggtcc 1440
                  ·        1400         ·        1420         ·        1440

     ------------------------------------------------------------

1441 agccgggtgcacacgctcagctcaaccgccagggcgtacacctggctgtacgccgccgcc 1500
                  ·        1460         ·        1480         ·        1500

     ------------------------------------------------------------

1501 agcagccccgacatcgccgccccaggggtctctagacctcgagtccggggagaacggtgg 1560
                  ·        1520         ·        1540         ·        1560

     ------------------------------------------------------------

1561 ccagacggcgcttgcgtctgcccccggagccctgccctcctccacccagcagcagcccgg 1620
                  ·        1580         ·        1600         ·        1620

     ------------------------------------------------------------

1621 ccgaggcctgcgacgcggtgctgaccggctcggccacgctgataaagttgtcctgggctg 1680
                  ·        1640         ·        1660         ·        1680

     ------------------------------------------------------------

1681 ccccgggcccaccccacactccctccagaaagtcccgagcggcctccgccgtccactcta 1740
                  ·        1700         ·        1720         ·        1740

     ------------------------------------------------------------

1741 tcccgctggaggcaatggtcgccagggtttctaggacgctgtccgccaggacggagaagc 1800
                  ·        1760         ·        1780         ·        1800

     ------------------------------------------------------------

1801 ggcccaataagtactccgcgtcgtccctagtcagcgaggcgcatgcctcgcccatggcat 1860
                  ·        1820         ·        1840         ·        1860

     ------------------------------------------------------------

1861 ccacaaggttgcacaccacatcaaacacacagtcttcctcctgtttttgtgatataatgg 1920
                  ·        1880         ·        1900         ·        1920

     ------------------------------------------------------------

1921 cctccaggccagccctgatgttctcaatctcatatgtggtcgcggcttgggtccggcgct 1980
                  ·        1940         ·        1960         ·        1980

     ------------------------------------------------------------

1981 tcacggtcaaccctagggtgggggtggcaaagacaaacttcttccgcatggaagagcccc 2040
                  ·        2000         ·        2020         ·        2040

     ------------------------------------------------------------

2041 cggcctgcttgcgcagcccagccccgggggcctgcagcaggttcctgtccacgccccggc 2100
                  ·        2060         ·        2080         ·        2100

     ------------------------------------------------------------

2101 ccataaagtatcccaggttcccggcctggaatatctggttgttgccgttgacccccgtgt 2160
                  ·        2120         ·        2140         ·        2160
```

```
     ------------------------------------------------------------

2161 acttgttgatggtcactggcagcgtgacaaccggacgggccttgcagacctggctaagac 2220
               ·       2180          ·       2200          ·       2220

     ------------------------------------------------------------

2221 agtctgtggccgcgcagaccaccgtggtcgcagtaagggaggaggtggcctccgcgtagg 2280
               ·       2240          ·       2260          ·       2280

     ------------------------------------------------------------

2281 ccgctgccgactccaccgcccgcgtgcccagtacgtgggggtagtcacgggcgggcaccg 2340
               ·       2300          ·       2320          ·       2340

   1 ----------------------atgctatcag----------------------------- 10
                          xxxxx|||||
2341 actgcgtcctcggcaccagtccctgaatcaggctgatgtagaactgggtctggccgcacg 2400
               ·       2360          ·       2380          ·       2400

     ------------------------------------------------------------

2401 ccttcaggatggcgttgttgagcctctgcttggcgtaagtgaccaggttgccaggcacca 2460
               ·       2420          ·       2440          ·       2460

     ------------------------------------------------------------

2461 catctatgacgttgctctcttcgtgggcccgggagcccccgtccacaaagagggccaggt 2520
               ·       2480          ·       2500          ·       2520

     ------------------------------------------------------------

2521 cagagtactcctccgcgctggccccgctggggacagggaccgagcgccgcctggaaaagt 2580
               ·       2540          ·       2560          ·       2580

     ------------------------------------------------------------

2581 tgtgccacaggtacaggcttgagagcttagtgtccgggaatagggtcttgtggtaggtgt 2640
               ·       2600          ·       2620          ·       2640

     ------------------------------------------------------------

2641 tgaggaatttcatgtagggcccgttgatgatgtagttctccctcctggtagtggacttga 2700
               ·       2660          ·       2680          ·       2700

     ------------------------------------------------------------

2701 tgaagctgttctggagggcggcattctcccccgtgaagaccaccctgttcttgatcttga 2760
               ·       2720          ·       2740          ·       2760

     ------------------------------------------------------------

2761 tgttcctggggcacagcatcagcaccttggacatgcgcacaggcagccgccggccgtaca 2820
               ·       2780          ·       2800          ·       2820

     ------------------------------------------------------------

2821 cccggccctgcagggccgcgtccaggtctggcaggtcgcaggtgggctccccatgcacca 2880
               ·       2840          ·       2860          ·       2880

     ------------------------------------------------------------

2881 ccttggcctccttggccgtgaggacccccttgtcgatggccaggctcctaaagttggtgc 2940
               ·       2900          ·       2920          ·       2940
```

```
     ------------------------------------------------------------

2941 acagcgtctggtagtgacccttagccactctgggggctctggccaagcccggggttgt 3000
               .       2960          .       2980          .       3000

     ------------------------------------------------------------

3001 cattctcatagcacatacagatgggcagggagatgtcctgcaggatggtcagcagtgagc 3060
               .       3020          .       3040          .       3060

     ------------------------------------------------------------

3061 ggtaaaacagctgggtgaagatggggcaggcgggctgcgcaaaggggttgcacgagtact 3120
               .       3080          .       3100          .       3120

     ------------------------------------------------------------

3121 gcatcacgtggtagcagctcttgaccaggtccttgtaggtgatgttgttcttggccatgc 3180
               .       3140          .       3160          .       3180

     ------------------------------------------------------------

3181 tgttcataaactggaccacttcggcgtccaccgccgcatccacgtccttgaacatcttga 3240
               .       3200          .       3220          .       3240

     ------------------------------------------------------------

3241 caaagtcacgcgggccatggggctccttctctagctttccttcagcgtctatgcccagcc 3300
               .       3260          .       3280          .       3300

     ------------------------------------------------------------

3301 gagacagccgctccagcaggttctggttcagctgccagtaggtgtagcggggctcgtcgt 3360
               .       3320          .       3340          .       3360

     ------------------------------------------------------------

3361 ccggccgctgcccgtcgtcctccttatcgatgaagttgagaaagttgcccaaaaagtccg 3420
               .       3380          .       3400          .       3420

     ------------------------------------------------------------

3421 tctcgttgtaggagcccgaggcccccgagatcacataggggtccctccgctgcgtggaca 3480
               .       3440          .       3460          .       3480

     ------------------------------------------------------------

3481 tgacgggggggaagcggtccctcagcctaaagaagagcgtgttcaggcacacggccgggg 3540
               .       3500          .       3520          .       3540

     ------------------------------------------------------------

3541 cccggccctcgcagagcgagcacatgggactggcggccgcccccgccacgtagctgcccg 3600
               .       3560          .       3580          .       3600

     ------------------------------------------------------------

3601 tctccggcaccggggtcagagagctcttctgtccctggcaaaactgcaggtagtaggcat 3660
               .       3620          .       3640          .       3660

     ------------------------------------------------------------

3661 agcgggcaagaaggttgggcgagaaggaggccgcatagaccaggtgctccacagcgtagt 3720
               .       3680          .       3700          .       3720
```

```
     ------------------------------------------------------------

3721 ttcccggaccgttggttccggtcacgtctggcccaccccagcccgagaagcagggtcggc 3780
               ·       3740          ·       3760          ·       3780

     ------------------------------------------------------------

3781 ggcaggggtcccaggtcccctcctgcagggtccccaggccgtgggtcatgtagaaactgt 3840
               ·       3800          ·       3820          ·       3840

     ------------------------------------------------------------

3841 taaagagactctccttgccctgaccggttgacttcgagacccccgagacgtagaggacgg 3900
               ·       3860          ·       3880          ·       3900

  11 ------------------------------------------------------------gta 13
                                                                   |||
3901 aattggtggcaaagatctgcgtggacacgtgggggggccaggctggcattatatcggtgta 3960
               ·       3920          ·       3940          ·       3960

  14 acgc-------------------------------------------------------- 17
     ||||
3961 acgcagccacacgggcctctggaccctcacagtcggcaaacaggggccacgagtcgtagt 4020
               ·       3980          ·       4000          ·       4020

     ------------------------------------------------------------

4021 tgaggctggccggggtctcgtgcgaggcctccagcatggcgggtgcgtagctcaccgcca 4080
               ·       4040          ·       4060          ·       4080

     ------------------------------------------------------------

4081 gctcgcatgccgcgctgtccacaatcattaaggctcccgagtccgggtgactgatggttg 4140
               ·       4100          ·       4120          ·       4140

     ------------------------------------------------------------

4141 aggctgggaactccttgaggggggccaccttggccaccttggcctggtcctgcaggctct 4200
               ·       4160          ·       4180          ·       4200

     ------------------------------------------------------------

4201 gcttctccagcagctccaccagcttgcccacccgtcggacgcgcagcgcctgcgccagcc 4260
               ·       4220          ·       4240          ·       4260

     ------------------------------------------------------------

4261 cggtgtacagcgcctcgtgcatgcagcggctgaggtccgagttgtaaaactggcggagct 4320
               ·       4280          ·       4300          ·       4320

     ------------------------------------------------------------

4321 ggggcacgccctctgggaacacctccttgtcgtagagcgggaccctaacgctcgcagact 4380
               ·       4340          ·       4360          ·       4380

     ------------------------------------------------------------

4381 gccccaccgctacctcctgttttaacgatggaatggccaccaggtttccgctgtagagtc 4440
               ·       4400          ·       4420          ·       4440

     ------------------------------------------------------------

4441 gctccttgaaggcctcggttattgccaccgcccccaggtaggcagagggatctagccctt 4500
               ·       4460          ·       4480          ·       4500
```

```
                                                                        
4501 cggggaagaagtcccccggcttggagctttccctcggtagggcgctgtaggcgtcgtacc 4560
                4520          4540          4560

                                                                        
4561 caaacacctccctggtctcgccacagagggcctcgagacccggcccctcaaagatggggg 4620
                4580          4600          4620

                                                                        
4621 gaaccatatgggcattgtggaacacgtagatgtccctgtgataggaggtagcgcgtagga 4680
                4640          4660          4680

                                                                        
4681 gcccgcagttggggtcgggcctcctgtgcagagccttgacattgatgctgaagcccggct 4740
                4700          4720          4740

                                                                        
4741 ccacggtgatgccgcaaaggagcggcaccgtcaggcacctgtggcccgcgtagccggtcc 4800
                4760          4780          4800

                                                                        
4801 ccagtgtggccacctccctaagagggtaggtggccaggggtaaaagtagatgtagccgc 4860
                4820          4840          4860

                                                                        
4861 acggacccggctggctctggctgcccagattatcctcgctagtctgtgcaccctgcatga 4920
                4880          4900          4920

                                                                        
4921 tgcccaaggtatcgccccggcctcccagtcccacattaaatgttacactttactcatcac 4980
                4940          4960          4980

                                                                        
4981 gcaacacccactgtttattcatttacaaagatttcaggaagtcagtcaggctggccaggg 5040
                5000          5020          5040

                                                                        
5041 cccacgtcacggggaactgacgtctcagcgatcttggcatgccgcccagcctcgcaaacc 5100
                5060          5080          5100

                                                                        
5101 agagtctgcgatagagggccaggtagtgggcgattgcccccagcacgaaggcggcgctct 5160
                5120          5140          5160

                                                                        
5161 tgtggtcatccaggtagtttcgcaccgcaaacaccactgtgtagcacagcaccaccctga 5220
                5180          5200          5220

                                                                        
5221 gccgcgaccagtagtcgtagtggtcgttgtacactgcgcgcaggacgctgatgatgagcc 5280
                5240          5260          5280
```

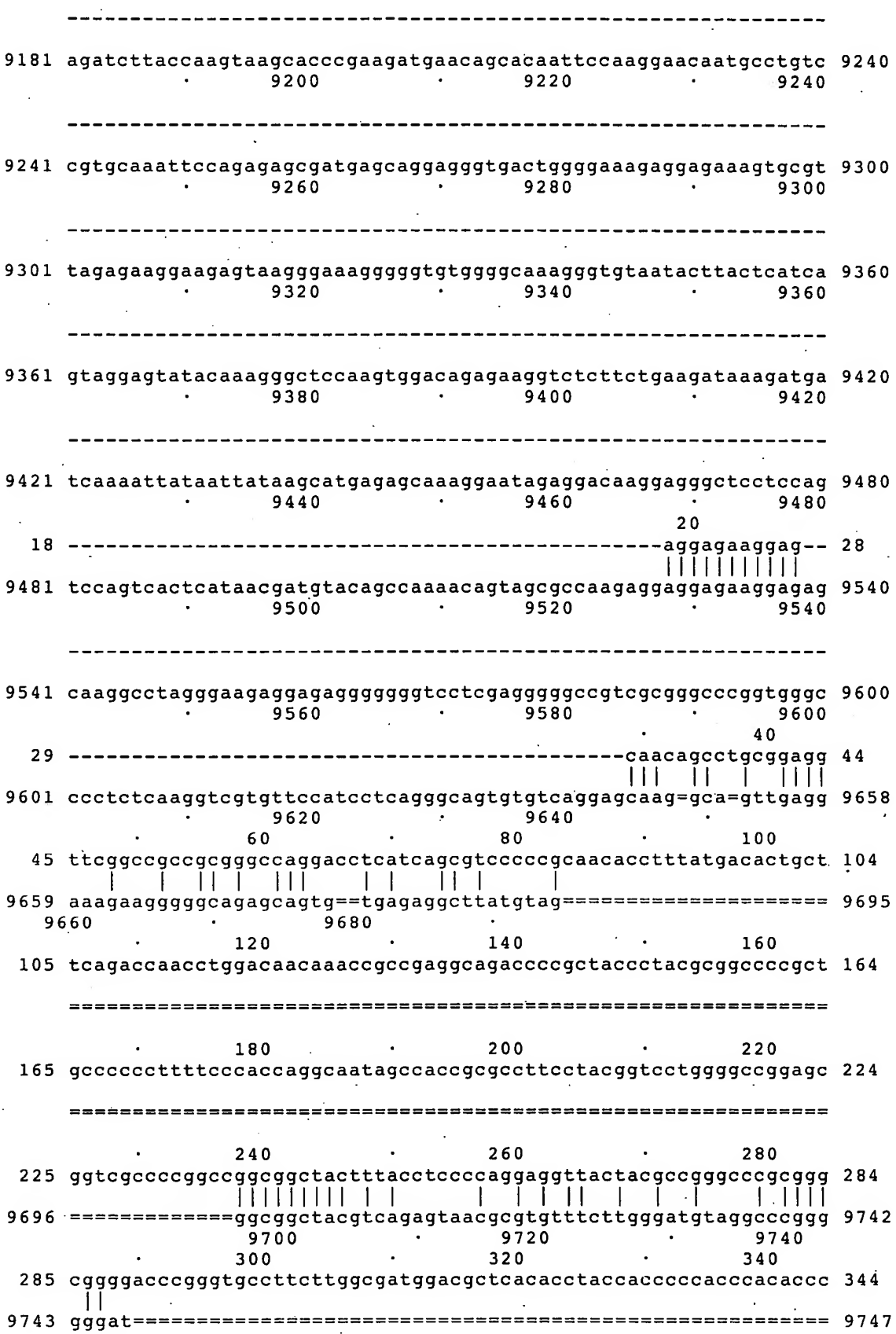
```
5281 gtacgtgcgtgtctttgcccccgatgtcggctgtcctgcaggccagctccgcgtacagct 5340
                 5300          5320          5340

5341 tcctatccttcctcagggaggccttgatgagccggcagaggaccagggctggcaaaggca 5400
                 5360          5380          5400

5401 ggtctttctcatcccgggtgaacaccgcgtacatggccctgaacatgaggtagctggact 5460
                 5420          5440          5460

5461 cagccaccttgtcgtccggcggcgagggcgcgacccacgcctcgaccggggtcctcacaa 5520
                 5480          5500          5520

5521 acacagaatctgtagacttggctggcctcatggtctcgtcaggccagctcacgggcttca 5580
                 5540          5560          5580

5581 ggcttatatgataaaatgggcgtggcagaatagtataagacgcgaggcctgggtgaggag 5640
                 5600          5620          5640

5641 agtccagagcaatggccaggttcatcgctcagctcctcctgttggcctcctgtgtggccg 5700
                 5660          5680          5700

5701 ccggccaggctgtcaccgctttcttgggtgagcgagtcaccctgacctcctactggagga 5760
                 5720          5740          5760

5761 gggtgagcctcggtccagagattgaggtcagctggtttaaactgggcccaggagaggagc 5820
                 5780          5800          5820

5821 aggtgcttattgggcgcatgcaccacgatgtcatctttatagagtggcctttcaggggct 5880
                 5840          5860          5880

5881 tctttgatatccacagaagtgccaacaccttcttttttagtagtcaccgctgccaacatct 5940
                 5900          5920          5940

5941 cccatgacggcaactacctgtgccgcatgaaactgggcgagaccgaggtcaccaagcagg 6000
                 5960          5980          6000

6001 aacacctgagcgtggtgaagcctctaacgctgtctgtccactccgaaaggtctcagttcc 6060
                 6020          6040          6060
```

```
6061 cagacttctctgtccttactgtgacatgcaccgtgaatgcatttccccatccccacgtcc 6120
                6080                6100                6120

6121 agtggctcatgcccgagggcgtggagcccgcaccaactgcggcaaatggcggtgttatga 6180
                6140                6160                6180

6181 aggaaaaggatgggagcctctctgttgctgttgacctgtcacttcccaagccctggcacc 6240
                6200                6220                6240

6241 tgccagtgacctgcgttgggaaaaatgacaaggaggaagcccacggggtttatgtttctg 6300
                6260                6280                6300

6301 gatacttgtcgcaataaacgcacttgcctatttcaccttgttttagtgtggcattggggg 6360
                6320                6340                6360

6361 ggtggcattgcgggtggatagcctcgcgactcgtgggaaaatgggcggaagggcaccgtg 6420
                6380                6400                6420

6421 ggaaaatagttccaggtgacagcagcagtgtgtgaagattgtcacagctgctggtttgga 6480
                6440                6460                6480

6481 gaaaacgggggtgggcggtgatcagggagaacaattccccggggacacctgcacgagacc 6540
                6500                6520                6540

6541 cctgggctctcaggaactccgcccaggtcttgccaattggggtgatcctgtagcgccgcg 6600
                6560                6580                6600

6601 gtttcagcatcacaggttattttgcctgaagcttgctggggcgtaaatccctctcgcctt 6660
                6620                6640                6660

6661 gtttctcagagagcatttcaggccggttttgcagtcgctgctgcagctatggggtcccta 6720
                6680                6700                6720

6721 gaaatggtgccaatgggcgcgggtccccctagccccggcggggatccggatgggtacgat 6780
                6740                6760                6780

6781 ggcggaaacaactcccaatatccatctgcttctggctcttctgggaacacccccacccca 6840
                6800                6820                6840
```

```
6841 ccgaacgatgaggaacgtgaatctaatgaagagcccccaccgccttatgaggacccatat 6900
                6860          6880          6900

6901 tggggcaatggcgaccgtcactcggactatcaaccactaggaacccaagatcaaagtctg 6960
                6920          6940          6960

6961 tacttgggattgcaacacgacgggaatgacgggctccctccccctccctactctccacgg 7020
                6980          7000          7020

7021 gatgactcatctcaacacatatacgaagaagcgggcagaggaaggtaagagtgccatcta 7080
                7040          7060          7080

7081 tctgtacttttatttattgcatcacaagtcacatcaataataagggcgccatctagcggg 7140
                7100          7120          7140

7141 agatgttatccacaccatcccaattcacatctcagggacaacaggtcaaagttctttgtt 7200
                7160          7180          7200

7201 gacacccccagcgctggctccaggggggtggaagcgttggatgcagtcctccgcatcgggg 7260
                7220          7240          7260

7261 cggacgcctcctcccaacgcgtttctgcggatcagtcgctggctggtgggcatcggagtc 7320
                7280          7300          7320

7321 ggtgggcggtcctccacggggacacgctccttcttggccttgttctttgacctttttggac 7380
                7340          7360          7380

7381 attcttctgaaggaacggcggagagtagcgtagaatccagccagtggtctacccggtcgc 7440
                7400          7420          7440

7441 atggtggcttcttagatgaggagcaggcataaaagtccaaacaggacacagagtaccacc 7500
                7460          7480          7500

7501 aggagtagtcttagtctgctgacgtctgggtcctcggggcaggggtggctaggcctggtc 7560
                7520          7540          7560

7561 tccgtagaagagccgggcaggccgcaggcagaggactgctgctctagcaaagcacgctcc 7620
                7580          7600          7620
```

```
7621 aggacgtgtaccatctcgagagtgaggcacagctgttttcgtggacttttatacagtaag 7680
                7640          7660          7680

7681 gacaaggaaagaaggccagaggaatgtggaaagatgagcgaggacaggtgtggaggtttt 7740
                7700          7720          7740

7741 gggctagctcttagtttctgggtgtgagagagggattaaagtgcttatgcgcaaagaatg 7800
                7760          7780          7800

7801 tgtcaacaacaggtgttcctgcctctgctggcatgagttaggtgtggcttgggctgaatc 7860
                7820          7840          7860

7861 caaatgtgtattggcacaagatggaaagcaaagttgctggagttactgggtgggagacag 7920
                7880          7900          7920

7921 ggatgtatgtggtcccccgctggtatgccagtaccctgtggaagtaaggggcctcatctg 7980
                7940          7960          7980

7981 cctggtagttgtgttgtgcagaggtctgatgtgtgtaggaggggtgggttcaacgcaggg 8040
                8000          8020          8040

8041 gcgttggtggcggagtctggcaacgcccgggtccttgctacctgtgtggtgtgttaaggg 8100
                8060          8080          8100

8101 ctgggtaaaggtgtctgccaattctcgcatgtcctcctttccccttgttttgaaatagaa 8160
                8120          8140          8160

8161 tatgaatgtggcttttcagcctagacagacagtgtggctaagggagtgtgtgccagttaa 8220
                8180          8200          8220

8221 ggtgattagctaaggcattcccagtaaatggagggagagtcagtcaggcaagcctatgac 8280
                8240          8260          8280

8281 atggtaatgcctagaagtaaagaaaggttagtcatagtagcttagctgaactgggccgtg 8340
                8300          8320          8340

8341 ggggtcgtcatcatctccaccggaaccagaagaacccaaaagcagcgtaggaaggtgtgg 8400
                8360          8380          8400
```

```
8401 atcaccgccgccatggccggaatcatgactatgaccgccgcctccgtctgtcatcaaagg 8460
                8420          8440          8460

8461 cgggccctggtcacctcctttgttttcaacctcttccgtcaattgtggagggcctccatc 8520
                8480          8500          8520

8521 atttccagcagagtcgctagggctatgaggcagcgggtcatgtgggccattgtcatcagt 8580
                8540          8560          8580

8581 gttgtcagggtcctgtgggccattgtcatcagtgttgtcagggtcctgaggcagcgggtc 8640
                8600          8620          8640

8641 atgtgggccattgtcatcagtgttgtcagggtcctgtgggccattgtcatcagtgttgtc 8700
                8660          8680          8700

8701 agggtcctgtgggccattgtcaggaccacctccaggtgcgcctaggttttgagagcagag 8760
                8720          8740          8760

8761 tgggggtccgtcgccggctccactcacgagcaggtggtgtctgccctcgttggagttaga 8820
                8780          8800          8820

8821 gtcagattcatggccagaatcatcggtagcttgttgagggtgcgggagggagtcatcgtg 8880
                8840          8860          8880

8881 gtggtgttcatcactgtgtcgttgtccatggtaatacatccagattaaaatcgccagaaa 8940
                8900          8920          8940

8941 caggaggagccaaaggagatcaaccaatagagtccaccagttttgttgtagatagagagc 9000
                8960          8980          9000

9001 aataatgagcaggatgaggtctaggaagaaggctaggaagaaggccaaaagctgccagat 9060
                9020          9040          9060

9061 ggtggcaccaagtcgccagagcatctccaataagtagatccagatacctaagactgcgtt 9120
                9080          9100          9120

9121 gaaaaaagagtgttagggttggaaaagtgggggtgtggtaaataattcctagggaatgtt 9180
                9140          9160          9180
```


9181 agatcttaccaagtaagcacccgaagatgaacagcacaattccaaggaacaatgcctgtc 9240
9241 cgtgcaaattccagagagcgatgagcaggagggtgactggggaaagaggagaaagtgcgt 9300
9301 tagagaaggaagagtaagggaaaggggggtgtggggcaaagggtgtaatacttactcatca 9360
9361 gtaggagtatacaaagggctccaagtggacagagaaggtctcttctgaagataaagatga 9420
9421 tcaaaattataattataagcatgagagcaaaggaatagaggacaaggagggctcctccag 9480
18 ----------aggagaaggag-- 28
9481 tccagtcactcataacgatgtacagccaaaacagtagcgccaagaggaggagaaggagag 9540
9541 caaggcctagggaagaggagaggggggggtcctcgaggggggccgtcgcgggcccggtgggc 9600
29 ----------caacagcctgcggagg 44
9601 ccctctcaaggtcgtgttccatcctcagggcagtgtgtcaggagcaag=gca=gttgagg 9658
45 ttcggccgccgcgggccaggacctcatcagcgtcccccgcaacacctttatgacactgct 104
9659 aaagaaggggggcagagcagtg==tgagaggcttatgtag========== 9695
105 tcagaccaacctggacaacaaaccgccgaggcagaccccgctaccctacgcggccccgct 164
165 gccccccttttcccaccaggcaatagccaccgcgccttcctacggtcctggggccggagc 224
225 ggtcgccccggccggcggctactttacctccccaggaggttactacgccgggcccgcggg 284
9696 ==========ggcggctacgtcagagtaacgcgtgtttcttgggatgtaggcccggg 9742
285 cggggacccgggtgccttcttggcgatggacgctcacacctaccacccccacccacaccc 344
9743 gggat========== 9747

```
          .       360         .       380         .       400
  345 ccctccggcctactttggcttgccgggcctctttggccccctccaccgtgcctccttac 404

      ============================================================

          .       420
  405 tacggattcccacttgcggg---------------------------------------- 424
                  |||||||
 9748 =============ttgcggggtctgccggaggcagtacgggtacagatttcccgaaagcg 9794
                    .      9760         .      9780         .

      ------------------------------------------------------------

 9795 gcggtgtgtgtgtgcatgtaagcgtagaaaggggaagtagaaagcgtgtgtttgtgttag 9854
        9800         .      9820         .      9840         .
           .       440         .       460         .       480
  425 cagactacgtccccg-ctccctcgcgatccaacaagcgg-aaaagagaccccgaggagga 482
       | |    |||||||     |  ||  |   | | ||||      ||  |   ||||  |
 9855 aaaagcgggtccccgggggcaagctgtggga=atgcggtggccaagtgcaacagga=aa 9912
        9860         .      9880         .      9900         .
           .       500         .       520         .       540
  483 tgaagaaggcgggggctat-tcccgggggaggacgccaccctctaccgcaaggacatag 541
      || | | |   | | ||| ||     |||||||
 9913 tggaaaggcagtgcggcaatcagaaggggag============================ 9944
          9920         .      9940
           .       560         .       580         .       600
  542 cgggcctctccaagagtgtgaatgagttacagcacacgctacaggccctgcgccgggaga 601

      ============================================================

           .       620         .       640         .       660
  602 cgctgtcctacggccacaccggagtcggatactgcccccagcagggcccctgctacaccc 661

      ============================================================

           .       680         .       700         .       720
  662 actcggggccttacggatttcagcctcatcaaagctacgaagtgcccagatacgtccctc 721

      ============================================================

           .       740         .       760         .       780
  722 atccgcccccaccaccaacttctcaccaggcagctcaggcgcagcctccacccccgggca 781
                           |     ||      |  |  || ||         | |  ||
 9945 ======================tgcgtagtgttgtgggaagcggcagtgtaatctgcaca 9982
                             .      9960         .      9980
           .       800         .       820         .
  782 cacaggcccccgaagcccactgtgtggccgagtccacgatccctgaggcgggag------ 835
       | |||| | .  |   || ||  || |||  | |||   |   |   |||||||||
 9983 aagaggcgcggggcgcgcaacgt=tgggag=gtcgttg==gcggcaggcgggaggccgtg 10038
           .      10000         .       10020         .

      ------------------------------------------------------------

10039 ctttaggggggttcaggtgaggcaaggctgtggggtaaccgtaggggaggcgggtgaggc 10098
   10040         .      10060         .      10080         .
                                                      840
  836 ----------------------------------------------cagccgggaa-ctc 848
                                                     |   |||||  ||
10099 ggctaagagggctaagggtcggcgggtgacgaagcagcagacggcggatatgggaatttc 10158
   10100         .      10120         .      10140         .
      .        860         .       880         .        900
  849 tggaccccg-ggaggacaccaaccctca-gcagcccac-caccgag-ggc--ca-cca-c 900
       | |    | || |||  |   |    | |  |    |    |||| | |   |  ||
10159 agaatgaggtggcggattcaggcgaaaagggtgtgggctgtgcgagtgtcatgaggcagg 10218
   10160         .      10180         .      10200         .
```

```
                920         940         960
  901 cgcggaaagaaactggtgcaggcctctgcgtccggagtggctcagtctaaggagcccacc 960
      |||||||||xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
10219 cgcggaaagtcgctgcggcttgctggggcatgggggggccgcgcattcctggaaaaagtgg 10278
   10220          10240         10260
                980         1000        1020
  961 acccccaaggccaagtctgtgtcagcccacctcaagtccatcttttgcgaggaattgctg 1020
      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
10279 agggggcgtggccttcccccgcggcccccccagccccccccgcacagagcggcgctacggcg 10338
   10280          10300         10320

 1021 aataaacgcgtggcttga---------------------------------------- 1038
      xxxxxxxxxxxxxxxxxx
10339 ggcgggcggcgggggggtcggggtccgcgggctccgggggctgcgggcggtggatggcggc 10398
   10340          10360         10380

      ------------------------------------------------------------

10399 ggacgttccggggatcgggggggtcgggggggcgccgcgcgggcgcagccatgcgtgaccg 10458
   10400          10420         10440

      ------------------------------------------------------------

10459 tgatgaggggggcagggtcgcaggggggtgtgtctggtgggggcgggagcgggggggcggcgc 10518
   10460          10480         10500

      ------------------------------------------------------------

10519 gggagcctgcacgccgttggagggtagaatgacaggggggcggggacagagaggcggtcgc 10578
   10520          10540         10560

      ------------------------------------------------------------

10579 gcccccggccgcgccagccaagcccccaaggggggcggggagcgggcaatggagcgtgac 10638
   10580          10600         10620

      ------------------------------------------------------------

10639 gaagggccccagggctgaccccggcaaacgtgacccggggctccggggtgacccagccaa 10698
   10640          10660         10680

      ------------------------------------------------------------

10699 gcgtgaccaaggggcccgtgggtgacacaggcaaccctgacaaaggcccccccaggaaaga 10758
   10700          10720         10740

      ------------------------------------------------------------

10759 cccccgggggggcatcgggggggtggggcatgggggggccgcgcattcctggaaaaagtggag 10818
   10760          10780         10800

      ------------------------------------------------------------

10819 ggggcgtggccttcccccgcggcccccccagccccccccgcacagagcggcgctacggcgggg 10878
   10820          10840         10860

      ------------------------------------------------------------

10879 cgggcggcgggggggtcggggtccgcgggctccgggggctgcgggcggtggatggcggcgg 10938
   10880          10900         10920

      ------------------------------------------------------------

10939 acgttccggggatcgggggggtcgggggggcgccgcgcgggcgcagccatgcgtgaccgtg 10998
   10940          10960         10980
```

```
10999 atgaggggggcagggtcgcaggggggtgtgtctggtgggggcgggagcgggggggcggcgcgg 11058
   11000          .      11020          .      11040          .

11059 gagcctgcacgccgttggagggtagaatgacaggggggcggggacagagaggcggtcgcgc 11118
   11060          .      11080          .      11100          .

11119 ccccggccgcgccagccaagcccccaagggggggcggggagcgggcaatggagcgtgacga 11178
   11120          .      11140          .      11160          .

11179 agggccccagggctgaccccggcaaacgtgacccggggctccggggtgacccagccaagc 11238
   11180          .      11200          .      11220          .

11239 gtgaccaaggggcccgtgggtgacacaggcaaccctgacaaaggccccccaggaaagacc 11298
   11240          .      11260          .      11280          .

11299 cccgtggggcatgggggggccgcgcattcctggaaaaagtggaggggggcgtggccttcccc 11358
   11300          .      11320          .      11340          .

11359 cgcggcccccccagcccccccgcacagagcggcgctacggcgggcgggcggcgggggggtcg 11418
   11360          .      11380          .      11400          .

11419 gggtccgcgggctccgggggctgcgggcggtggatggcggcggacgttccggggatcggg 11478
   11420          .      11440          .      11460          .

11479 ggggtcgggggggcgccgcgcgggcgcagccatgcgtgaccgtgatgaggggggcagggtcg 11538
   11480          .      11500          .      11520          .

11539 caggggggtgtgtctggtgggggcgggagcgggggggcggcgcgggagcctgcacgccgttg 11598
   11540          .      11560          .      11580          .

11599 gagggtagaatgacaggggggcggggacagagaggcggtcgcgcccccggccgcgccagcc 11658
   11600          .      11620          .      11640          .

11659 aagcccccaagggggggcggggagcgggcaatggagcgtgacgaagggccccagggctgac 11718
   11660          .      11680          .      11700          .

11719 cccggcaaacgtgacccggggctccggggtgacccagccaagcgtgaccaaggggcccgt 11778
   11720          .      11740          .      11760          .
```

```
      ------------------------------------------------------------

11779 gggtgacacaggcaaccctgacaaaggccccccaggaaagacccccgggggggcatcgggg 11838
   11780          ·      11800          ·      11820          ·

      ------------------------------------------------------------

11839 ggtggggcatgggggggccgcgcattcctggaaaaagtggaggggggcgtggccttcccccg 11898
   11840          ·      11860          ·      11880          ·

      ------------------------------------------------------------

11899 cggcccccccagcccccccgcacagagcggcgctacggcgggcgggcggcgggggggtcggg 11958
   11900          ·      11920          ·      11940          ·

      ------------------------------------------------------------

11959 gtccgcgggctccgggggctgcgggcggtggatggcggcggacgttccggggatcgggggg 12018
   11960          ·      11980          ·      12000          ·

      ------------------------------------------------------------

12019 ggtcgggggggcgccgcgcgggcgcagccatgcgtgaccgtgatgaggggggcagggtcgca 12078
   12020          ·      12040          ·      12060          ·

      ------------------------------------------------------------

12079 gggggtgtgtctggtgggggcgggagcgggggggcggcgcgggagcctgcacgccgttgga 12138
   12080          ·      12100          ·      12120          ·

      ------------------------------------------------------------

12139 gggtagaatgacaggggggcggggacagagaggcggtcgcgccccccggccgcgccagccaa 12198
   12140          ·      12160          ·      12180          ·

      ------------------------------------------------------------

12199 gcccccaagggggggcggggagcgggcaatggagcgtgacgaagggccccagggctgaccc 12258
   12200          ·      12220          ·      12240          ·

      ------------------------------------------------------------

12259 cggcaaacgtgacccggggctccggggtgacccagccaagcgtgaccaaggggcccgtgg 12318
   12260          ·      12280          ·      12300          ·

      ------------------------------------------------------------

12319 gtgacacaggcaaccctgacaaaggccccccaggaaagacccccgggggggcatcgggggg 12378
   12320          ·      12340          ·      12360          ·

      ------------------------------------------------------------

12379 ggtgttggcgggggcatgggggggtcggatttcgcccttattgccctgtttagaattc   12436
   12380          ·      12400          ·      12420          ·

  % Identity =   1.6 (210/12958)

///
```

9310-13DVCTDV SEQ ID NO 3.xdna x Bankier et al. EcoRI Dhet fragment complement.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

DNA sequence 12436 bp gaattctaaaca ... ctttgagaattc linear

Method: Blocks (Martinez)
Layout: Standard
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Translation: Off

Alignment 10. Comparison of nucleotide sequence of SEQ ID NO:3 with the complement of the nucleotide sequence of Fig. 2 of Bankier et al.

```
    ------------------------------------------------------------

  1 gaattctaaacagggcaataagggcgaaatccgacccccccatgcccccgccaacacccc 60
              .         20          .         40          .         60

    ------------------------------------------------------------

 61 ccccgatgcccccccgggggtctttcctgggggggcctttgtcagggttgcctgtgtcaccc 120
              .         80          .        100          .        120

    ------------------------------------------------------------

121 acgggcccccttggtcacgcttggctgggtcaccccggagccccgggtcacgtttgccggg 180
              .        140          .        160          .        180

    ------------------------------------------------------------

181 gtcagccctggggcccttcgtcacgctccattgcccgctccccgccccccttgggggctt 240
              .        200          .        220          .        240

    ------------------------------------------------------------

241 ggctggcgcggccgggggcgcgaccgcctctctgtccccgccccctgtcattctaccctc 300
              .        260          .        280          .        300

    ------------------------------------------------------------

301 caacggcgtgcaggctcccgcgccgcccccccgctcccgcccccaccagacacacccccctg 360
              .        320          .        340          .        360

    ------------------------------------------------------------

361 cgaccctgcccccctcatcacggtcacgcatggctgcgcccgcgcggcgcccccccgacccc 420
              .        380          .        400          .        420

    ------------------------------------------------------------

421 cccgatccccggaacgtccgccgccatccaccgcccgcagcccccggagcccgcggaccc 480
              .        440          .        460          .        480

    ------------------------------------------------------------

481 cgacccccccgccgcccgcccgccgtagcgccgctctgtgcgggggggctgggggggccgcg 540
              .        500          .        520          .        540

    ------------------------------------------------------------

541 ggggaaggccacgcccccctccacttttttccaggaatgcgcggcccccccatgccccacccc 600
              .        560          .        580          .        600
```

```
601 ccgatgcccccgggggtctttcctgggggccttgtcagggttgcctgtgtcacccac 660
          .       620         .       640         .       660

661 gggccccttggtcacgcttggctgggtcaccccggagccccgggtcacgtttgccggggt 720
          .       680         .       700         .       720

721 cagccctggggcccttcgtcacgctccattgcccgctccccgccccccttgggggcttgg 780
          .       740         .       760         .       780

781 ctggcgcggccgggggcgcgaccgcctctctgtccccgccccctgtcattctaccctcca 840
          .       800         .       820         .       840

841 acggcgtgcaggctcccgcgccgcccccgctcccgcccccaccagacacacccccctgcg 900
          .       860         .       880         .       900

901 accctgccccctcatcacggtcacgcatggctgcgcccgcgcggcgcccccgaccccc 960
          .       920         .       940         .       960

961 cgatccccggaacgtccgccgccatccaccgcccgcagcccccggagcccgcggaccccg 1020
          .       980         .      1000         .      1020

1021 acccccgccgcccgcccgccgtagcgccgctctgtgcgggggggctgggggccgcggg 1080
          .      1040         .      1060         .      1080

1081 ggaaggccacgcccccctccacttttttccaggaatgcgcggcccccccatgccccacggggg 1140
          .      1100         .      1120         .      1140

1141 tctttcctgggggggcctttgtcagggttgcctgtgtcacccacgggcccccttggtcacgc 1200
          .      1160         .      1180         .      1200

1201 ttggctgggtcaccccggagccccgggtcacgtttgccggggtcagccctggggcccttc 1260
          .      1220         .      1240         .      1260

1261 gtcacgctccattgcccgctccccgcccccctttgggggcttggctggcgcggccgggggc 1320
          .      1280         .      1300         .      1320

1321 gcgaccgcctctctgtccccgcccccctgtcattctaccctccaacggcgtgcaggctccc 1380
          .      1340         .      1360         .      1380
```

```
     ------------------------------------------------------------

1381 gcgccgccccccgctcccgccccaccagacacacccccctgcgaccctgccccctcatca 1440
               .       1400          .       1420          .       1440

     ------------------------------------------------------------

1441 cggtcacgcatggctgcgcccgcgcggcgcccccccgaccccccccgatccccggaacgtcc 1500
               .       1460          .       1480          .       1500

     ------------------------------------------------------------

1501 gccgccatccaccgcccgcagcccccggagcccgcggaccccgaccccccgccgcccgcc 1560
               .       1520          .       1540          .       1560

     ------------------------------------------------------------

1561 cgccgtagcgccgctctgtgcgggggggctgggggccgcgggggaaggccacgccccct 1620
               .       1580          .       1600          .       1620

     ------------------------------------------------------------

1621 ccacttttccaggaatgcgcggcccccatgccccaccccccgatgccccccgggggtc 1680
               .       1640          .       1660          .       1680

     ------------------------------------------------------------

1681 tttcctgggggggcctttgtcagggttgcctgtgtcacccacgggccccttggtcacgctt 1740
               .       1700          .       1720          .       1740

     ------------------------------------------------------------

1741 ggctgggtcaccccggagccccgggtcacgtttgccggggtcagccctggggcccttcgt 1800
               .       1760          .       1780          .       1800

     ------------------------------------------------------------

1801 cacgctccattgcccgctccccgcccccctttgggggcttggctggcgcggccgggggcgc 1860
               .       1820          .       1840          .       1860

     ------------------------------------------------------------

1861 gaccgcctctctgtccccgcccccctgtcattctaccctccaacggcgtgcaggctcccgc 1920
               .       1880          .       1900          .       1920

     ------------------------------------------------------------

1921 gccgcccccgctcccgccccaccagacacacccccctgcgaccctgccccctcatcacg 1980
               .       1940          .       1960          .       1980

     ------------------------------------------------------------

1981 gtcacgcatggctgcgcccgcgcggcgcccccccgaccccccccgatccccggaacgtccgc 2040
               .       2000          .       2020          .       2040

     ------------------------------------------------------------

2041 cgccatccaccgcccgcagcccccggagcccgcggaccccgaccccccgccgcccgcccg 2100
               .       2060          .       2080          .       2100

     ------------------------------------------------------------

2101 ccgtagcgccgctctgtgcgggggggctgggggccgcgggggaaggccacgcccccctcc 2160
               .       2120          .       2140          .       2160
```

```
     ------------------------------------------------------------

2161 acttttccaggaatgcgcggccccccatgccccagcaagccgcagcgactttccgcgcc 2220
               ·       2180          ·       2200          ·       2220

     ------------------------------------------------------------

2221 tgcctcatgacactcgcacagcccacacccttttcgcctgaatccgccacctcattctga 2280
               ·       2240          ·       2260          ·       2280

     ------------------------------------------------------------

2281 aattcccatatccgccgtctgctgcttcgtcacccgccgacccttagccctcttagccgc 2340
               ·       2300          ·       2320          ·       2340

     ------------------------------------------------------------

2341 ctcacccgcctcccctacggttaccccacagccttgcctcacctgaaccccccctaaagca 2400
               ·       2360          ·       2380          ·       2400

     ------------------------------------------------------------

2401 cggcctcccgcctgccgccaacgacctcccaacgttgcgcgccccgcgcctctttgtgca 2460
               ·       2420          ·       2440          ·       2460

     ------------------------------------------------------------

2461 gattacactgccgcttcccacaacactacgcactcccccttctgattgccgcactgcctt 2520
               ·       2480          ·       2500          ·       2520

     ------------------------------------------------------------

2521 tccatttcctgttgcacttggccaccgcattcccacagcttgcccccggggacccgctt 2580
               ·       2540          ·       2560          ·       2580

     ------------------------------------------------------------

2581 ttctaacacaaacacacgctttctacttcccctttctacgcttacatgcacacacacacc 2640
               ·       2600          ·       2620          ·       2640

     ------------------------------------------------------------

2641 gccgctttcgggaaatctgtacccgtactgcctccggcagaccccgcaaatccccccggg 2700
               ·       2660          ·       2680          ·       2700

     ------------------------------------------------------------

2701 cctacatcccaagaaacacgcgttactctgacgtagccgccctacataagcctctcacac 2760
               ·       2720          ·       2740          ·       2760

     ------------------------------------------------------------

2761 tgctctgccccctttctttcctcaactgccttgctcctgacacactgccctgaggatggaa 2820
               ·       2780          ·       2800          ·       2820

     ------------------------------------------------------------

2821 cacgaccttgagaggggcccaccgggcccgcgacggcccccctcgaggacccccccctctcc 2880
               ·       2840          ·       2860          ·       2880

     ------------------------------------------------------------

2881 tcttccctaggccttgctctccttctcctcctcttggcgctactgttttggctgtacatc 2940
               ·       2900          ·       2920          ·       2940
```

```
     ----------------------------------------------------------------

2941 gttatgagtgactggactggaggagccctccttgtcctctattcctttgctctcatgctt 3000
               .      2960          .      2980          .      3000

     ----------------------------------------------------------------

3001 ataattataattttgatcatctttatcttcagaagagaccttctctgtccacttggagcc 3060
               .      3020          .      3040          .      3060

     ----------------------------------------------------------------

3061 ctttgtatactcctactgatgagtaagtattacacccttgccccacacccccttccct 3120
               .      3080          .      3100          .      3120
                    .        20
   1 ---atgctatcaggtaacgcagg---------------------------------------- 20
        xxxxxxxxxxx|||||||xx
3121 tactcttccttctctaacgcactttctcctctttccccagtcaccctcctgctcatcgct 3180
               .      3140          .      3160          .      3180

     ----------------------------------------------------------------

3181 ctctggaatttgcacggacaggcattgttccttggaattgtgctgttcatcttcgggtgc 3240
               .      3200          .      3220          .      3240

     ----------------------------------------------------------------

3241 ttacttggtaagatctaacattccctaggaattatttaccacacccccacttttccaacc 3300
               .      3260          .      3280          .      3300

     ----------------------------------------------------------------

3301 ctaacactctttttcaacgcagtcttaggtatctggatctacttattggagatgctctg 3360
               .      3320          .      3340          .      3360

     ----------------------------------------------------------------

3361 gcgacttggtgccaccatctggcagcttttggccttcttcctagccttcttcctagacct 3420
               .      3380          .      3400          .      3420

     ----------------------------------------------------------------

3421 catcctgctcattattgctctctatctacaacaaaactggtggactctattggttgatct 3480
               .      3440          .      3460          .      3480

     ----------------------------------------------------------------

3481 cctttggctcctcctgtttctggcgattttaatctggatgtattaccatggacaacgaca 3540
               .      3500          .      3520          .      3540

     ----------------------------------------------------------------

3541 cagtgatgaacaccaccacgatgactccctcccgcaccctcaacaagctaccgatgattc 3600
               .      3560          .      3580          .      3600

     ----------------------------------------------------------------

3601 tggccatgaatctgactctaactccaacgagggcagacaccacctgctcgtgagtggagc 3660
               .      3620          .      3640          .      3660

     ----------------------------------------------------------------

3661 cggcgacggacccccactctgctctcaaaacctaggcgcacctggaggtggtcctgacaa 3720
               .      3680          .      3700          .      3720
```

```
     ----------------------------------------------------------------

3721 tggcccacaggaccctgacaacactgatgacaatggcccacaggaccctgacaacactga 3780
                .       3740          .       3760          .       3780

     ----------------------------------------------------------------

3781 tgacaatggcccacatgacccgctgcctcaggaccctgacaacactgatgacaatggccc 3840
                .       3800          .       3820          .       3840

     ----------------------------------------------------------------

3841 acaggaccctgacaacactgatgacaatggcccacatgacccgctgcctcatagccctag 3900
                .       3860          .       3880          .       3900

     ----------------------------------------------------------------

3901 cgactctgctggaaatgatggaggccctccacaattgacggaagaggttgaaaacaaagg 3960
                .       3920          .       3940          .       3960

     ----------------------------------------------------------------

3961 aggtgaccagggcccgcctttgatgacagacggaggcggcggtcatagtcatgattccgg 4020
                .       3980          .       4000          .       4020

     ----------------------------------------------------------------

4021 ccatggcggcggtgatccacaccttcctacgctgcttttgggttcttctggttccggtgg 4080
                .       4040          .       4060          .       4080

     ----------------------------------------------------------------

4081 agatgatgacgacccccacggcccagttcagctaagctactatgactaacctttctttac 4140
                .       4100          .       4120          .       4140

     ----------------------------------------------------------------

4141 ttctaggcattaccatgtcataggcttgcctgactgactctccctccatttactgggaat 4200
                .       4160          .       4180          .       4200

     ----------------------------------------------------------------

4201 gccttagctaatcaccttaactggcacacactcccttagccacactgtctgtctaggctg 4260
                .       4220          .       4240          .       4260

     ----------------------------------------------------------------

4261 aaaagccacattcatattctatttcaaaacaaggggaaaggaggacatgcgagaattggc 4320
                .       4280          .       4300          .       4320

     ----------------------------------------------------------------

4321 agacacctttacccagcccttaacacaccacacaggtagcaaggacccgggcgttgccag 4380
                .       4340          .       4360          .       4380

     ----------------------------------------------------------------

4381 actccgccaccaacgcccctgcgttgaacccacccctcctacacacatcagacctctgca 4440
                .       4400          .       4420          .       4440

     ----------------------------------------------------------------

4441 caacacaactaccaggcagatgaggccccttacttccacagggtactggcataccagcgg 4500
                .       4460          .       4480          .       4500
```

```
     ------------------------------------------------------------

4501 gggaccacatacatccctgtctcccacccagtaactccagcaactttgctttccatcttg 4560
               ·    4520          ·       4540          ·       4560

     ------------------------------------------------------------

4561 tgccaatacacatttggattcagcccaagccacacctaactcatgccagcagaggcagga 4620
               ·       4580          ·       4600          ·       4620

     ------------------------------------------------------------

4621 acacctgttgttgacacattctttgcgcataagcactttaatccctctctcacacccaga 4680
               ·       4640          ·       4660          ·       4680

     ------------------------------------------------------------

4681 aactaagagctagcccaaaacctccacacctgtcctcgctcatctttccacattcctctg 4740
               ·       4700          ·       4720          ·       4740

     ------------------------------------------------------------

4741 gccttctttccttgtccttactgtataaaagtccacgaaaacagctgtgcctcactctcg 4800
               ·       4760          ·       4780          ·       4800

     ------------------------------------------------------------

4801 agatggtacacgtcctggagcgtgctttgctagagcagcagtcctctgcctgcggcctgc 4860
               ·       4820          ·       4840          ·       4860

     ------------------------------------------------------------

4861 ccggctcttctacggagaccaggcctagccacccctgccccgaggacccagacgtcagca 4920
               ·       4880          ·       4900          ·       4920

     ------------------------------------------------------------

4921 gactaagactactcctggtggtactctgtgtcctgtttggacttttatgcctgctcctca 4980
               ·       4940          ·       4960          ·       4980

     ------------------------------------------------------------

4981 tctaagaagccaccatgcgaccgggtagaccactggctggattctacgctactctccgcc 5040
               ·       5000          ·       5020          ·       5040

  21 ---------------------------------------------agaaggagc--------- 29
                                                    |||||||||
5041 gttccttcagaagaatgtccaaaaggtcaaagaacaaggccaagaaggagcgtgtccccg 5100
               ·       5060          ·       5080          ·       5100

     ------------------------------------------------------------

5101 tggaggaccgcccaccgactccgatgcccaccagccagcgactgatccgcagaaacgcgt 5160
               ·       5120          ·       5140          ·       5160

     ------------------------------------------------------------

5161 tgggaggaggcgtccgccccgatgcggaggactgcatccaacgcttccacccccctggagc 5220
               ·       5180          ·       5200          ·       5220

     ------------------------------------------------------------

5221 cagcgctgggggtgtcaacaaagaactttgacctgttgtccctgagatgtgaattgggat 5280
               ·       5240          ·       5260          ·       5280
```

```
     ------------------------------------------------------------

5281 ggtgtggataacatctcccgctagatggcgcccttattattgatgtgacttgtgatgcaa 5340
               .       5300          .       5320          .       5340

     ------------------------------------------------------------

5341 taaataaaagtacagatagatggcactcttaccttcctctgcccgcttcttcgtatatgt 5400
               .       5360          .       5380          .       5400

     ------------------------------------------------------------

5401 gttgagatgagtcatcccgtggagagtagggaggggggagggagcccgtcattcccgtcgt 5460
               .       5420          .       5440          .       5460

     ------------------------------------------------------------

5461 gttgcaatcccaagtacagactttgatcttgggttcctagtggttgatagtccgagtgac 5520
               .       5480          .       5500          .       5520

     ------------------------------------------------------------

5521 ggtcgccattgccccaatatgggtcctcataaggcggtgggggctcttcattagattcac 5580
               .       5540          .       5560          .       5580

     ------------------------------------------------------------

5581 gttcctcatcgttcggtggggtgggggtgttcccagaagagccagaagcagatggatatt 5640
               .       5600          .       5620          .       5640

     ------------------------------------------------------------

5641 gggagttgtttccgccatcgtacccatccggatccccgccggggctaggggacccgcgc 5700
               .       5660          .       5680          .       5700

     ------------------------------------------------------------

5701 ccattggcaccatttctagggaccccatagctgcagcagcgactgcaaaaccggcctgaa 5760
               .       5720          .       5740          .       5760

     ------------------------------------------------------------

5761 atgctctctgagaaacaaggcgagagggatttacgccccagcaagcttcaggcaaaataa 5820
               .       5780          .       5800          .       5820

     ------------------------------------------------------------

5821 cctgtgatgctgaaaccgcggcgctacaggatcaccccaattggcaagacctgggcggag 5880
               .       5840          .       5860          .       5880

     ------------------------------------------------------------

5881 ttcctgagagcccaggggtctcgtgcaggtgtccccggggaattgttctccctgatcacc 5940
               .       5900          .       5920          .       5940

     ------------------------------------------------------------

5941 gcccacccccgttttctccaaaccagcagctgtgacaatcttcacacactgctgctgtca 6000
               .       5960          .       5980          .       6000

     ------------------------------------------------------------

6001 cctggaactattttcccacggtgcccttccgcccattttcccacgagtcgcgaggctatc 6060
               .       6020          .       6040          .       6060
```

```
     ------------------------------------------------------------

6061 cacccgcaatgccacccccccaatgccacactaaaacaaggtgaaataggcaagtgcgtt 6120
               ·       6080          ·       6100          ·       6120

     ------------------------------------------------------------

6121 tattgcgacaagtatccagaaacataaaccccgtgggcttcctccttgtcatttttccca 6180
               ·       6140          ·       6160          ·       6180

  30 ------------------------------------------------aacag----- 34
                                                      |||||
6181 acgcaggtcactggcaggtgccagggcttgggaagtgacaggtcaacagcaacagagagg 6240
               ·       6200          ·       6220          ·       6240

     ------------------------------------------------------------

6241 ctcccatccttttccttcataacaccgccatttgccgcagttggtgcgggctccacgccc 6300
               ·       6260          ·       6280          ·       6300

     ------------------------------------------------------------

6301 tcgggcatgagccactggacgtggggatggggaaatgcattcacggtgcatgtcacagta 6360
               ·       6320          ·       6340          ·       6360

     ------------------------------------------------------------

6361 aggacagagaagtctgggaactgagacctttcggagtggacagacagcgttagaggcttc 6420
               ·       6380          ·       6400          ·       6420

     ------------------------------------------------------------

6421 accacgctcaggtgttcctgcttggtgacctcggtctcgcccagtttcatgcggcacagg 6480
               ·       6440          ·       6460          ·       6480

     ------------------------------------------------------------

6481 tagttgccgtcatgggagatgttggcagcggtgactactaaaaagaaggtgttggcactt 6540
               ·       6500          ·       6520          ·       6540

     ------------------------------------------------------------

6541 ctgtggatatcaaagaagcccctgaaaggccactctataaagatgacatcgtggtgcatg 6600
               ·       6560          ·       6580          ·       6600

     ------------------------------------------------------------

6601 cgcccaataagcacctgctcctctcctgggcccagtttaaaccagctgacctcaatctct 6660
               ·       6620          ·       6640          ·       6660

     ------------------------------------------------------------

6661 ggaccgaggctcaccctcctccagtaggaggtcagggtgactcgctcacccaagaaagcg 6720
               ·       6680          ·       6700          ·       6720

  35 -------cctg------------------------------------------------- 38
             ||||
6721 gtgacagcctggccggcggccacacaggaggccaacaggaggagctgagcgatgaacctg 6780
               ·       6740          ·       6760          ·       6780

     ------------------------------------------------------------

6781 gccattgctctggactctcctcacccaggcctcgcgtcttatactattctgccacgccca 6840
               ·       6800          ·       6820          ·       6840
```

```
     ----------------------------------------------------------

6841 ttttatcatataagcctgaagcccgtgagctggcctgacgagaccatgaggccagccaag 6900
               ·       6860          ·       6880          ·       6900

     ----------------------------------------------------------

6901 tctacagattctgtgtttgtgaggaccccggtcgaggcgtgggtcgcgccctcgccgccg 6960
               ·       6920          ·       6940          ·       6960

     ----------------------------------------------------------

6961 gacgacaaggtggctgagtccagctacctcatgttcagggccatgtacgcggtgttcacc 7020
               ·       6980          ·       7000          ·       7020

     ----------------------------------------------------------

7021 cgggatgagaaagacctgcctttgccagccctggtcctctgccggctcatcaaggcctcc 7080
               ·       7040          ·       7060          ·       7080
                                                            40
  39 ----------------------------------------------------------cggaggt 45
                                                               ||| ||
7081 ctgaggaaggataggaagctgtacgcggagctggcctgcaggacagccgacatcgggggc 7140
               ·       7100          ·       7120          ·       7140
          ·         60          ·         80          ·
  46 tcggccgccgcgggccaggacctcatcagcgtcccccgcaacacctttatgaca------ 99
     | | |     |  | |  | ||||||||||||xxxxxxxxxxxxxxxxxxxx
7141 aaagacacgcacgtacggctcatcatcagcgtcctgcgcgcagtgtacaacgaccactac 7200
               ·       7160          ·       7180          ·       7200

     ----------------------------------------------------------

7201 gactactggtcgcggctcagggtggtgctgtgctacacagtggtgtttgcggtgcgaaac 7260
               ·       7220          ·       7240          ·       7260

     ----------------------------------------------------------

7261 tacctggatgaccacaagagcgccgccttcgtgctgggggcaatcgcccactacctggcc 7320
               ·       7280          ·       7300          ·       7320

     ----------------------------------------------------------

7321 ctctatcgcagactctggtttgcgaggctgggcggcatgccaagatcgctgagacgtcag 7380
               ·       7340          ·       7360          ·       7380

     ----------------------------------------------------------

7381 ttccccgtgacgtgggccctggccagcctgactgacttcctgaaatctttgtaaatgaat 7440
               ·       7400          ·       7420          ·       7440

     ----------------------------------------------------------

7441 aaacagtgggtgttgcgtgatgagtaaagtgtaacatttaatgtgggactgggaggccgg 7500
               ·       7460          ·       7480          ·       7500

     ----------------------------------------------------------

7501 ggcgataccttgggcatcatgcagggtgcacagactagcgaggataatctgggcagccag 7560
               ·       7520          ·       7540          ·       7560

     ----------------------------------------------------------

7561 agccagccgggtccgtgcggctacatctacttttacccccctggccacctaccctcttagg 7620
               ·       7580          ·       7600          ·       7620
```

```
     ------------------------------------------------------------

7621 gaggtggccacactggggaccggctacgcgggccacaggtgcctgacggtgccgctcctt 7680
          .       7640          .       7660          .       7680

     ------------------------------------------------------------

7681 tgcggcatcaccgtggagccgggcttcagcatcaatgtcaaggctctgcacaggaggccc 7740
          .       7700          .       7720          .       7740

     ------------------------------------------------------------

7741 gaccccaactgcgggctcctacgcgctacctcctatcacagggacatctacgtgttccac 7800
          .       7760          .       7780          .       7800

     ------------------------------------------------------------

7801 aatgcccatatggttccccccatctttgaggggccgggtctcgaggccctctgtggcgag 7860
          .       7820          .       7840          .       7860

     ------------------------------------------------------------

7861 accagggaggtgtttgggtacgacgcctacagcgccctaccgagggaaagctccaagccg 7920
          .       7880          .       7900          .       7920

     ------------------------------------------------------------

7921 ggggacttcttccccgaagggctagatccctctgcctacctgggggcggtggcaataacc 7980
          .       7940          .       7960          .       7980

     ------------------------------------------------------------

7981 gaggccttcaaggagcgactctacagcggaaacctggtggccattccatcgttaaaacag 8040
          .       8000          .       8020          .       8040

     ------------------------------------------------------------

8041 gaggtagcggtggggcagtctgcgagcgttagggtcccgctctacgacaaggaggtgttc 8100
          .       8060          .       8080          .       8100

     ------------------------------------------------------------

8101 ccagagggcgtgccccagctccgccagttttacaactcggacctcagccgctgcatgcac 8160
          .       8120          .       8140          .       8160

     ------------------------------------------------------------

8161 gaggcgctgtacaccgggctggcgcaggcgctgcgcgtccgacgggtgggcaagctggtg 8220
          .       8180          .       8200          .       8220

     ------------------------------------------------------------

8221 gagctgctggagaagcagagcctgcaggaccaggccaaggtggccaaggtggccccctc 8280
          .       8240          .       8260          .       8280

     ------------------------------------------------------------

8281 aaggagttcccagcctcaaccatcagtcacccggactcgggagccttaatgattgtggac 8340
          .       8300          .       8320          .       8340

     ------------------------------------------------------------

8341 agcgcggcatgcgagctggcggtgagctacgcacccgccatgctggaggcctcgcacgag 8400
          .       8360          .       8380          .       8400
```

```
     ------------------------------------------------------------

8401 accccggccagcctcaactacgactcgtggcccctgtttgccgactgtgagggtccagag 8460
                .       8420          .       8440          .       8460

     ------------------------------------------------------------

8461 gcccgtgtggctgcgttacaccgatataatgccagcctggccccccacgtgtccacgcag 8520
                .       8480          .       8500          .       8520

     ------------------------------------------------------------

8521 atctttgccaccaattccgtcctctacgtctcgggggtctcgaagtcaaccggtcagggc 8580
                .       8540          .       8560          .       8580

     ------------------------------------------------------------

8581 aaggagagtctctttaacagtttctacatgacccacggcctggggaccctgcaggagggg 8640
                .       8600          .       8620          .       8640
                          100
 100 -----------------------ctgcttc------------------------------ 106
                             |||||||
8641 acctgggacccctgccgccgaccctgcttctcgggctggggtgggccagacgtgaccgga 8700
                .       8660          .       8680          .       8700

 107 ------------------------------------------------------------agacc 111
                                                                  | ||
8701 accaacggtccgggaaactacgctgtggagcacctggtctatgcggcctccttctcgccc 8760
                .       8720          .       8740          .       8760

 112 aacct------------------------------------------------------- 116
     |||||
8761 aaccttcttgcccgctatgcctactacctgcagttttgccagggacagaagagctctctg 8820
                .       8780          .       8800          .       8820

     ------------------------------------------------------------

8821 accccggtgccggagacgggcagctacgtggcgggggcggccgccagtcccatgtgctcg 8880
                .       8840          .       8860          .       8880

     ------------------------------------------------------------

8881 ctctgcgagggccgggccccggccgtgtgcctgaacacgctcttctttaggctgagggac 8940
                .       8900          .       8920          .       8940

     ------------------------------------------------------------

8941 cgcttcccccccgtcatgtccacgcagcggagggacccctatgtgatctcgggggcctcg 9000
                .       8960          .       8980          .       9000

 117 ------------------------------------------------------------g 117

9001 ggctcctacaacgagacggacttttgggcaactttctcaacttcatcgataaggaggac 9060
                .       9020          .       9040          .       9060
     120        .          140
 118 gacaacaaaccgcc-gaggcaga-ccccgctac---------------------------- 148
     |||    | | ||| || |  || |||||||||
9061 gacgggcagcggccggacgacgagccccgctacacctactggcagctgaaccagaacctg 9120
                .       9080          .       9100          .       9120

     ------------------------------------------------------------

9121 ctggagcggctgtctcggctgggcatagacgctgaaggaaagctagagaaggagccccat 9180
                .       9140          .       9160          .       9180
```

```
     ------------------------------------------------------------

9181 ggcccgcgtgactttgtcaagatgttcaaggacgtggatgcggcggtggacgccgaagtg 9240
               ·      9200          ·      9220          ·      9240

     ------------------------------------------------------------

9241 gtccagtttatgaacagcatggccaagaacaacatcacctacaaggacctggtcaagagc 9300
               ·      9260          ·      9280          ·      9300

     ------------------------------------------------------------

9301 tgctaccacgtgatgcagtactcgtgcaacccctttgcgcagcccgcctgccccatcttc 9360
               ·      9320          ·      9340          ·      9360

     ------------------------------------------------------------

9361 acccagctgttttaccgctcactgctgaccatcctgcaggacatctccctgcccatctgt 9420
               ·      9380          ·      9400          ·      9420

     ------------------------------------------------------------

9421 atgtgctatgagaatgacaaccccgggcttggccagagccccccagagtggctaaagggt 9480
               ·      9440          ·      9460          ·      9480

     ------------------------------------------------------------

9481 cactaccagacgctgtgcaccaactttaggagcctggccatcgacaaggggtcctcacg 9540
               ·      9500          ·      9520          ·      9540

     ------------------------------------------------------------

9541 gccaaggaggccaaggtggtgcatggggagcccacctgcgacctgccagacctggacgcg 9600
               ·      9560          ·      9580          ·      9600

     ------------------------------------------------------------

9601 gccctgcagggccgggtgtacggccggcggctgcctgtgcgcatgtccaaggtgctgatg 9660
               ·      9620          ·      9640          ·      9660

     ------------------------------------------------------------

9661 ctgtgccccaggaacatcaagatcaagaacagggtggtcttcacgggggagaatgccgcc 9720
               ·      9680          ·      9700          ·      9720

     ------------------------------------------------------------

9721 ctccagaacagcttcatcaagtccactaccaggagggagaactacatcatcaacgggccc 9780
               ·      9740          ·      9760          ·      9780

     ------------------------------------------------------------

9781 tacatgaaattcctcaacacctaccacaagaccctattcccggacactaagctctcaagc 9840
               ·      9800          ·      9820          ·      9840

     ------------------------------------------------------------

9841 ctgtacctgtggcacaacttttccaggcggcgctcggtccctgtccccagcggggccagc 9900
               ·      9860          ·      9880          ·      9900

     ------------------------------------------------------------

9901 gcggaggagtactctgacctggccctctttgtggacgggggctcccgggcccacgaagag 9960
               ·      9920          ·      9940          ·      9960
```

```
      ------------------------------------------------------------

 9961 agcaacgtcatagatgtggtgcctggcaacctggtcacttacgccaagcagaggctcaac 10020
                ·      9980          ·     10000          ·     10020

      ------------------------------------------------------------

10021 aacgccatcctgaaggcgtgcggccagacccagttctacatcagcctgattcagggactg 10080
                ·     10040          ·     10060          ·     10080

      ------------------------------------------------------------

10081 gtgccgaggacgcagtcggtgcccgcccgtgactacccccacgtactgggcacgcgggcg 10140
                ·     10100          ·     10120          ·     10140

  149 -----------------cctacgcgg---------------------------------- 157
                       |||||||||
10141 gtggagtcggcagcggcctacgcggaggccacctcctcccttactgcgaccacggtggtc 10200
                ·     10160          ·     10180          ·     10200
                                                          160
  158 ----------------------------------------------------ccccgctgccc 168
                                                           | |||||||x
10201 tgcgcggccacagactgtcttagccaggtctgcaaggcccgtccggttgtcacgctgcca 10260
                ·     10220          ·     10240          ·     10260
        ·       180
  169 ccctttcccaccagg------------------------------------------- 184
      xxxxxxxxxxxxxxx
10261 gtgaccatcaacaagtacacgggggtcaacggcaacaaccagatattccaggccgggaac 10320
                ·     10280          ·     10300          ·     10320

      ------------------------------------------------------------

10321 ctgggatactttatgggccggggcgtggacaggaacctgctgcaggcccccggggctggg 10380
                ·     10340          ·     10360          ·     10380

      ------------------------------------------------------------

10381 ctgcgcaagcaggccgggggctcttccatgcggaagaagtttgtctttgccacccccacc 10440
                ·     10400          ·     10420          ·     10440

      ------------------------------------------------------------

10441 ctagggttgaccgtgaagcgccggacccaagccgcgaccacatatgagattgagaacatc 10500
                ·     10460          ·     10480          ·     10500

      ------------------------------------------------------------

10501 agggctggcctggaggccattatatcacaaaaacaggaggaagactgtgtgtttgatgtg 10560
                ·     10520          ·     10540          ·     10560

      ------------------------------------------------------------

10561 gtgtgcaaccttgtggatgccatgggcgaggcatgcgcctcgctgactagggacgacgcg 10620
                ·     10580          ·     10600          ·     10620

      ------------------------------------------------------------

10621 gagtacttattgggccgcttctccgtcctggcggacagcgtcctagaaaccctggcgacc 10680
                ·     10640          ·     10660          ·     10680

      ------------------------------------------------------------

10681 attgcctccagcgggatagagtggacggcggaggccgctcgggactttctggagggagtg 10740
                ·     10700          ·     10720          ·     10740
```

```
      ------------------------------------------------------------

10741 tggggtgggcccggggcagcccaggacaactttatcagcgtggccgagccggtcagcacc 10800
                .     10760         .     10780         .     10800

      ------------------------------------------------------------

10801 gcgtcgcaggcctcggccgggctgctgctgggtggaggagggcagggctccgggggcaga 10860
                .     10820         .     10840         .     10860

      ------------------------------------------------------------

10861 cgcaagcgccgtctggccaccgttctccccggactcgaggtctagagacccctggggcgg 10920
                .     10880         .     10900         .     10920

      ------------------------------------------------------------

10921 cgatgtcggggctgctggcggcggcgtacagccaggtgtacgccctggcggttgagctga 10980
                .     10940         .     10960         .     10980

      ------------------------------------------------------------

10981 gcgtgtgcacccggctggacccccggagtctggacgtggctgcggtggtgcgcaacgccg 11040
                .     11000         .     11020         .     11040

      ------------------------------------------------------------

11041 gcctgctggccgagctggaggccatcctccttccccgtttgagacggcagaatgaccgtg 11100
                .     11060         .     11080         .     11100

      ------------------------------------------------------------

11101 catgcagcgccctgtccctggagctggtgcacctgctagagaactcgagagaggcctctg 11160
                .     11120         .     11140         .     11160

      ------------------------------------------------------------

11161 ccgcgctgctcgcccctggtagaaagggtacccgggtcccgcctctccgtacccccctcag 11220
                .     11180         .     11200         .     11220

      ------------------------------------------------------------

11221 tcgcgtactctgtggagttttacgggggggcataaagtcgatgtaagtttgtgcctaataa 11280
                .     11240         .     11260         .     11280

  185 -----------------------------caatagc------------------------- 191
                                    |||||||
11281 atgacatagagattttaatgaagagaatcaatagcgtgttttattgcatgtctcacacca 11340
                .     11300         .     11320         .     11340

      ------------------------------------------------------------

11341 tggggctggagagcctggaacgggccctggatctgctgggccgctttcggggcgtaagtc 11400
                .     11360         .     11380         .     11400

      ------------------------------------------------------------

11401 ccatcccagacccgcgcctctacatcacctctgtgccctgctggcgctgtgtgggcgagc 11460
                .     11420         .     11440         .     11460

      ------------------------------------------------------------

11461 tgatggttctgcccaaccacggcaacccttccacggcagaggggacccacgtctcctgta 11520
                .     11480         .     11500         .     11520
```

```
      ------------------------------------------------------------

11521 accacctggcggtgccggtgaatccggagccggtctcgggactgtttgagaatgaagtcc 11580
                ·     11540         ·     11560         ·     11580

      ------------------------------------------------------------

11581 gccaggcggggctcgggcacctgttggaggctgaggagaaggcgaggccgggcggcccag 11640
                ·     11600         ·     11620         ·     11640
              200         ·       220
  192 --caccgcgccttcctacggtcctgg-ggccggag------------------------- 223
        |  ||||  |||   |  |  || ||||||||
11641 aggagggcgcggtcccgggcccggggcggccggaggcagaggggggcgaccagagcgctgg 11700
                ·     11660         ·     11680         ·     11700

      ------------------------------------------------------------

11701 acacctacaacgtcttctcgacagtgcccccggaggtggcggagctctcggagctcctct 11760
                ·     11720         ·     11740         ·     11760

      ------------------------------------------------------------

11761 attggaactctggcggccatgctatcggtgcaacggggcaggggggagggtggcggccatt 11820
                ·     11780         ·     11800         ·     11820

      ------------------------------------------------------------

11821 cccgcctctctgccctgtttgcccgggagcgtcgcctggccctggtgcgggggggcctgcg 11880
                ·     11840         ·     11860         ·     11880

      ------------------------------------------------------------

11881 aggaggcgctggcgggggcaaggctgactcacctgtttgacgccgtggctcccggggcca 11940
                ·     11900         ·     11920         ·     11940

  224 -----cgg-tc------------------------------------------------- 228
           ||| ||
11941 cggagcggctcttctgcggcggggtctacagctcctcgggcgacgcggtggaggcgctga 12000
                ·     11960         ·     11980         ·     12000

      ------------------------------------------------------------

12001 aggcggactgcgccgccgccttcacggcgcacccccagtaccgggccatcctgcaaaaga 12060
                ·     12020         ·     12040         ·     12060

      ------------------------------------------------------------

12061 ggaacgagctgtacacgcggctcaaccgagccatgcagcggttgggccgaggcgaggagg 12120
                ·     12080         ·     12100         ·     12120
                                  ·       240         ·       260
  229 -------------------------gccccggccggcggctactttacctccccaggagg 263
                                |||||||||||
12121 aggcgtcccgggagagcccggaggtgccccggccggc======================== 12157
                ·     12140         ·
              ·       280         ·       300         ·       320
  264 ttactacgccgggcccgcgggcggggacccgggtgccttcttggcgatggacgctcacac 323

      ============================================================

              ·       340         ·       360         ·       380
  324 ctaccacccccacccacacccccctccggcctactttggcttgccgggcctctttggccc 383

      ============================================================
```

```
            .       400         .       420         .       440
  384 ccctccaccgtgcctccttactacggattcccacttgcgggcagactacgtccccgctcc 443

      ============================================================

            .       460         .       480         .       500
  444 ctcgcgatccaacaagcggaaaagagaccccgaggaggatgaagaaggcgggggctatt 503
                                      || | | | |||   |||   |
12158 ================================tggggcacg=agagcccggcccgtccgg 12184
                                  12160         .     12180
            .       520         .       540         .       560
  504 cccgggggaggacgccaccctctaccgcaaggacatagcgggcctctccaagagtgtgaa 563
      | |      |||||   | ||| | ||||||||
12185 cgccctctcggacg===cgctcaagcgcaagga========================== 12214
            .       12200
            .       580         .       600         .       620
  564 tgagttacagcacacgctacaggccctgcgccgggagacgctgtcctacggccacaccgg 623
                      |||  |||||||| || |
12215 ==================gcagtacctgcgccaggtgg===================== 12234
                      12220         .
            .       640         .       660         .       680
  624 agtcggatactgcccccagcagggcccctgctacacccactcggggccttacggatttca 683

      ============================================================

            .       700         .       720         .       740
  684 gcctcatcaaagctacgaagtgcccagatacgtccctcatccgccccccaccaccaacttc 743

      ============================================================

            .       760         .       780         .       800
  744 tcaccaggcagctcaggcgcagcctccacccccgggcacacaggcccccgaagcccactg 803

      ============================================================

            .       820         .       840         .       860
  804 tgtggccgagtccacgatccctgaggcgggagcagccgggaactctggaccccgggagga 863

      ============================================================

            .       880         .       900         .       920
  864 caccaaccctcagcagcccaccaccgagggccaccaccgcggaaagaaactggtgcaggc 923
                      ||||||||||
12235 ==================ccaccgagggtct========================== 12247
                     12240
            .       940         .       960         .       980
  924 ctctgcgtccggagtggctcagtctaaggagcccaccaccccaaggccaagtctgtgtc 983
                                                  |||||||     |||
12248 ============================================ggccaagctgcagtc 12262
                                                  .     12260
            .         1000          .       1020        .
  984 a-gcccacctcaa--gtcc-atcttttgcgag-gaattgctgaataaacgcgtggcttga 1038
        |||   | |||  |  | |     ||   | ||     ||  |   ||  |||   ||
12263 ctgcctggcgcaacagagcgagaccctgaccgagaccctgtgcctgcgcgtctgggggga 12322
            .     12280         .     12300         .     12320

      ------------------------------------------------------------

12323 cgtggtctactgggagctggcccgcatgcgcaaccacttcctctacagacgggccttcgt 12382
            .     12340         .     12360         .     12380

      ------------------------------------------------------

12383 ctcgggtccctgggaggacaggcgcgccggcgagggtgccgcctttgagaattc       12436
            .     12400         .     12420         .

  % Identity =   1.8 (239/13074)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:33:28 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ*

Protein sequence 609 aa *GRRGVLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 11. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
              .         20          .         40          .         60
  1 MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
      RR        R      P +P      Q    +              +       Q  Y+
  1 *GRRGVLIGPLLRPGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGWARGSPGQL=YQRG 59
              .         20          .         40          .
              .         80          .        100          .        120
 61 VQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAASAGTGALASSAPST 120
        V     RA     +RA + G          S  T    +    AA +G  A A S
 60 RAGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYA 119
   60         .          80          .        100          .
              .         140          .        160          .
121 -AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*-- 177
     AV  S    +       + A    A   A   A +             +A
120 LAVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALSLELVHLLEN 179
  120         .         140          .        160          .

    ------------------------------------------------------------

180 SREASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFY 239
  180         .         200          .        220          .

    ------------------------------------------------------------

240 CMSHTMGLESLERALDLLGRFRGVSPIPDPRLYITSVPCWRCVGELMVLPNHGNPSTAEG 299
  240         .         260          .        280          .

    ------------------------------------------------------------

300 THVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAEEKARPGGPEEGAVPGPGRPEAEG 359
  300         .         320          .        340          .

    ------------------------------------------------------------

360 ATRALDTYNVFSTVPPEVAELSELLYWNSGGHAIGATGQGEGGGHSRLSALFARERRLAL 419
  360         .         380          .        400          .

    ------------------------------------------------------------

420 VRGACEEALAGARLTHLFDAVAPGATERLFCGGVYSSSGDAVEALKADCAAAFTAHPQYR 479
  420         .         440          .        460          .

    ------------------------------------------------------------

480 AILQKRNELYTRLNRAMQRLGRGEEEASRESPEVPRPAGAREPGPSGALSDALKRKEQYL 539
  480         .         500          .        520          .

    ------------------------------------------------------------

540 RQVATEGLAKLQSCLAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRRAFVSGPWEDR 599
  540         .         560          .        580          .
```

---

600 RAGEGAAFEN 609
600

% Identity = 4.9 (30/610) % Homology = 2.0 (12/610) % Total = 6.9 (42/610)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA*

Protein sequence 609 aa *GRRGVLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 12. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20                .         40          .
  1 MLSGNAGEGATAC-GGSAAAGQ-D--LIS--VPRNTFMTLLQTNLDNKP--PRQTPLPYA 52
          G     GG       D  L    V       L   ++      P Q
  1 *GRRGVLIGPLLRPGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGWARGSPGQLYQRGR 60
              .         20          .         40          .         60
        60          .         80          .         100         .
 53 APLPPFSHQAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPH 112
    A     +    A A     G       A     SP        G       LA      +
 61 AGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYAL 120
              .         80          .         100         .         120
        120           .         140         .         160         .
113 PHPPPAYFGL-PG-LFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPG 170
           L P  L             + L A  +P   R N R       E    L
121 AVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALSLELVHLLENS 180
              .         140         .         160         .         180
          180         .         200         .         220         .
171 EDATLYRKDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPH 230
     +A+            V  L+    A   E        V  C
181 REASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFYC 240
              .         200         .         220         .         240
          240         .         260         .         280
231 QSYEVPRYVPHPPPPPTSHQAAQAQPPPPG---TQAPEAHCVAESTI-PEAGAAGNS-GP 285
     S+ +                    +  P P    T  P   CV E  + P  G    + G
241 MSHTMGLESLERALDLLGRFRGVSPIPDPRLYITSVPCWRCVGELMVLPNHGNPSTAEGT 300
              .         260         .         280         .         300
        .         300         .         320         .         340
286 REDTNPQQ-PTTEGHHRGKKL--V-QASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLN 341
       N    P       G     V QA    + +++E   P       A       E
301 HVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAEEKARPGGPEEGAVPGPGRPEAEGA 360
              .         320         .         340         .         360

342 KRVA*------------------------------------------------------- 346
     R
361 TRALDTYNVFSTVPPEVAELSELLYWNSGGHAIGATGQGEGGGHSRLSALFARERRLALV 420
              .         380         .         400         .         420

    ------------------------------------------------------------

421 RGACEEALAGARLTHLFDAVAPGATERLFCGGVYSSSGDAVEALKADCAAAFTAHPQYRA 480
              .         440         .         460         .         480

    ------------------------------------------------------------

481 ILQKRNELYTRLNRAMQRLGRGEEEASRESPEVPRPAGAREPGPSGALSDALKRKEQYLR 540
              .         500         .         520         .         540

    ------------------------------------------------------------

541 QVATEGLAKLQSCLAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRRAFVSGPWEDRR 600
              .         560         .         580         .         600
```

---

601 AGEGAAFEN 609

% Identity = 9.7 (59/609) % Homology = 3.0 (18/609) % Total = 12.6 (77/609)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 609 aa *GRRGVLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 13. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 *GRRGVLIGPLLRPGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGWARGSPGQLYQRGR 60
                ·         20          ·         40          ·         60

      ------------------------------------------------------------

   61 AGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYAL 120
                ·         80          ·        100          ·        120

      ------------------------------------------------------------

  121 AVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALSLELVHLLENS 180
                ·        140          ·        160          ·        180

      ------------------------------------------------------------

  181 REASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFYC 240
                ·        200          ·        220          ·        240

      ------------------------------------------------------------

  241 MSHTMGLESLERALDLLGRFRGVSPIPDPRLYITSVPCWRCVGELMVLPNHGNPSTAEGT 300
                ·        260          ·        280          ·        300

      ------------------------------------------------------------

  301 HVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAEEKARPGGPEEGAVPGPGRPEAEGA 360
                ·        320          ·        340          ·        360

      ------------------------------------------------------------

  361 TRALDTYNVFSTVPPEVAELSELLYWNSGGHAIGATGQGEGGGHSRLSALFARERRLALV 420
                ·        380          ·        400          ·        420

      ------------------------------------------------------------

  421 RGACEEALAGARLTHLFDAVAPGATERLFCGGVYSSSGDAVEALKADCAAAFTAHPQYRA 480
                ·        440          ·        460          ·        480

      ------------------------------------------------------------

  481 ILQKRNELYTRLNRAMQRLGRGEEEASRESPEVPRPAGAREPGPSGALSDALKRKEQYLR 540
                ·        500          ·        520          ·        540

    1 -------------------------------------------AVDTGSGGGGQPHDT 15
                                                            G
  541 QVATEGLAKLQSCLAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRRAFVSGPWEDRR 600
                ·        560          ·        580          ·        600
```

```
          20
    16 APRGARKKQ                                                  24
       A  GA  +
   601 AGEGAAFEN                                                  609
```

% Identity = 0.7 (4/609) % Homology = 0.2 (1/609) % Total = 0.8 (5/609)

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:35:15 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 609 aa *GRRGVLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 14. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
     ------------------------------------------------------------

   1 *GRRGVLIGPLLRPGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGWARGSPGQLYQRGR 60
               .         20         .         40         .         60

     ------------------------------------------------------------

  61 AGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYAL 120
               .         80         .        100         .        120

     ------------------------------------------------------------

 121 AVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALSLELVHLLENS 180
               .        140         .        160         .        180

     ------------------------------------------------------------

 181 REASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFYC 240
               .        200         .        220         .        240

     ------------------------------------------------------------

 241 MSHTMGLESLERALDLLGRFRGVSPIPDPRLYITSVPCWRCVGELMVLPNHGNPSTAEGT 300
               .        260         .        280         .        300

     ------------------------------------------------------------

 301 HVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAEEKARPGGPEEGAVPGPGRPEAEGA 360
               .        320         .        340         .        360

     ------------------------------------------------------------

 361 TRALDTYNVFSTVPPEVAELSELLYWNSGGHAIGATGQGEGGGHSRLSALFARERRLALV 420
               .        380         .        400         .        420

     ------------------------------------------------------------

 421 RGACEEALAGARLTHLFDAVAPGATERLFCGGVYSSSGDAVEALKADCAAAFTAHPQYRA 480
               .        440         .        460         .        480

     ------------------------------------------------------------

 481 ILQKRNELYTRLNRAMQRLGRGEEEASRESPEVPRPAGAREPGPSGALSDALKRKEQYLR 540
               .        500         .        520         .        540
                                                         .         20
   1 ---------------------------------------STAVAQSATPSVSSSISSLRA 21
                                                           +  S     R
 541 QVATEGLAKLQSCLAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRRAFVSGPWEDRR 600
               .        560         .        580         .        600
```

```
    22 ATSGATAAA                                                    30
       A  GA
   601 AGEGAAFEN                                                   609
```

% Identity = 0.8 (5/609) % Homology = 0.2 (1/609) % Total = 1.0 (6/609)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ*

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV*

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 15. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                                                            
  1 MQGAQTSEDNLGSQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
              .         20          .         40          .         60

                                                            
 61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMVPPIFEGPGLEALCGETREVFG 120
              .         80          .        100          .        120

                                                            
121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
              .        140          .        160          .        180

                                                            
181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
              .        200          .        220          .        240

                                                            
241 SLQDQAKVAKVAPLKEFPASTISHPDSGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
              .        260          .        280          .        300

                                                            
301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
              .        320          .        340          .        360

                                                            
361 SFYMTHGLGTLQEGTWDPCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
              .        380          .        400          .        420

                                                            
421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
              .        440          .        460          .        480

                                                            
481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
              .        500          .        520          .        540

                                                            
541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
              .        560          .        580          .        600
```

```
     ------------------------------------------------------------

601  YSCNPFAQPACPIFTQLFYRSLLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTLC 660
               .        620          .        640          .        660

     ------------------------------------------------------------

661  TNFRSLAIDKGVLTAKEAKVVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720
               .        680          .        700          .        720

     ------------------------------------------------------------

721  KIKNRVVFTGENAALQNSFIKSTTRRENYIINGPYMKFLNTYHKTLFPDTKLSSLYLWHN 780
               .        740          .        760          .        780

     ------------------------------------------------------------

781  FSRRRSVPVPSGASAEEYSDLALFVDGGSRAHEESNVIDVVPGNLVTYAKQRLNNAILKA 840
               .        800          .        820          .        840

     ------------------------------------------------------------

841  CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900
               .        860          .        880          .        900

   1 ----------------------------------------------------MARRLPKP 8
                                                            xxxxxxxx
901  LSQVCKARPVVTLPVTINKYTGVNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960
               .        920          .        940          .        960
      .         20          .         40          .         60
   9 TLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEYVQRTFGVP 68
     xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
961  GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLVD 1020
               .        980          .       1000          .       1020
      .         80          .        100          .        120
  69 RRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAAASAGTGALASSAPSTAVAQSATP 128
     xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAARDFLEGVWGGPGA 1080
               .       1040          .       1060          .       1080
      .        140          .        160          .
 129 SVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*             177
     xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1081 AQDNFISVAEPVSTASQASAGLLLGGGGQGSGGRRKRRLATVLPGLEV*             1129
               .       1100          .       1120

 % Identity =   0.0 (0/1129)       % Homology =   0.0 (0/1129)       % Total =   0.0 (0/1129)

///
```

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA*

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV*

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 16. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
    ------------------------------------------------------------

  1 MQGAQTSEDNLGSQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
              .         20          .         40          .         60

    ------------------------------------------------------------

 61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMVPPIFEGPGLEALCGETREVFG 120
              .         80          .        100          .        120

    ------------------------------------------------------------

121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
              .        140          .        160          .        180

    ------------------------------------------------------------

181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
              .        200          .        220          .        240

    ------------------------------------------------------------

241 SLQDQAKVAKVAPLKEFPASTISHPDSGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
              .        260          .        280          .        300

    ------------------------------------------------------------

301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
              .        320          .        340          .        360

    ------------------------------------------------------------

361 SFYMTHGLGTLQEGTWDPCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
              .        380          .        400          .        420

    ------------------------------------------------------------

421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
              .        440          .        460          .        480

    ------------------------------------------------------------

481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
              .        500          .        520          .        540

    ------------------------------------------------------------

541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
              .        560          .        580          .        600
```

```
     ------------------------------------------------------------

   601 YSCNPFAQPACPIFTQLFYRSLLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTLC 660
                 .       620         .       640         .       660

     ------------------------------------------------------------

   661 TNFRSLAIDKGVLTAKEAKVVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720
                 .       680         .       700         .       720

     ------------------------------------------------------------

   721 KIKNRVVFTGENAALQNSFIKSTTRRENYIINGPYMKFLNTYHKTLFPDTKLSSLYLWHN 780
                 .       740         .       760         .       780
                   .        20         .        40         .
     1 ---MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPP 57
        xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   781 FSRRRSVPVPSGASAEEYSDLALFVDGGSRAHEESNVIDVVPGNLVTYAKQRLNNAILKA 840
                 .       800         .       820         .       840
       60          .        80         .       100         .
      58 FSHQAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPP 117
         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   841 CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900
                 .       860         .       880         .       900
       120         .       140         .       160         .
      118 AYFGLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYR 177
         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   901 LSQVCKARPVVTLPVTINKYTGVNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960
                 .       920         .       940         .       960
       180         .       200         .       220         .
      178 KDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPR 237
         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   961 GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLVD 1020
                 .       980         .      1000         .      1020
       240         .       260         .       280         .
      238 YVPHPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTE 297
         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAARDFLEGVWGGPGA 1080
                 .      1040         .      1060         .      1080
       300         .       320         .       340
      298 GHHRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*             346
         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  1081 AQDNFISVAEPVSTASQASAGLLLGGGGQGSGGRRKRRLATVLPGLEV*             1129
                 .      1100         .      1120

  % Identity =   0.0 (0/1129)      % Homology =   0.0 (0/1129)      % Total =   0.0 (0/1129)

///
```

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

```
Protein  sequence       24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence     1129  aa   MQGAQTSEDNLG ... RLATVLPGLEV*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 17. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 MQGAQTSEDNLGSQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
                ·         20          ·         40          ·         60

      ------------------------------------------------------------

   61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMVPPIFEGPGLEALCGETREVFG 120
                ·         80          ·        100          ·        120

      ------------------------------------------------------------

  121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
                ·        140          ·        160          ·        180

      ------------------------------------------------------------

  181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
                ·        200          ·        220          ·        240

      ------------------------------------------------------------

  241 SLQDQAKVAKVAPLKEFPASTISHPDSGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
                ·        260          ·        280          ·        300

      ------------------------------------------------------------

  301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
                ·        320          ·        340          ·        360

      ------------------------------------------------------------

  361 SFYMTHGLGTLQEGTWDPCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
                ·        380          ·        400          ·        420

      ------------------------------------------------------------

  421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
                ·        440          ·        460          ·        480

      ------------------------------------------------------------

  481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
                ·        500          ·        520          ·        540

      ------------------------------------------------------------

  541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
                ·        560          ·        580          ·        600
```

```
    ------------------------------------------------------------

601 YSCNPFAQPACPIFTQLFYRSLLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTLC 660
              .       620         .       640         .       660

    ------------------------------------------------------------

661 TNFRSLAIDKGVLTAKEAKVVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720
              .       680         .       700         .       720

    ------------------------------------------------------------

721 KIKNRVVFTGENAALQNSFIKSTTRRENYIINGPYMKFLNTYHKTLFPDTKLSSLYLWHN 780
              .       740         .       760         .       780

    ------------------------------------------------------------

781 FSRRRSVPVPSGASAEEYSDLALFVDGGSRAHEESNVIDVVPGNLVTYAKQRLNNAILKA 840
              .       800         .       820         .       840

    ------------------------------------------------------------

841 CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900
              .       860         .       880         .       900

    ------------------------------------------------------------

901 LSQVCKARPVVTLPVTINKYTGVNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960
              .       920         .       940         .       960

    ------------------------------------------------------------

961 GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLVD 1020
              .       980         .      1000         .      1020

    ------------------------------------------------------------

1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAARDFLEGVWGGPGA 1080
              .      1040         .      1060         .      1080
                                      .       20
   1 -------------------------AVDTGSGGGGQPHDTAPRGARKKQ             24
                               xxxxxxxxxxxxxxxxxxxxxxxx
1081 AQDNFISVAEPVSTASQASAGLLLGGGGQGSGGRRKRRLATVLPGLEV*             1129
              .      1100         .      1120

% Identity =   0.0 (0/1129)      % Homology =   0.0 (0/1129)      % Total =   0.0 (0/1129)
```

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV*

Method: Diagonals (BLOSUM62)
Layout: Standard
Block Length ≤: 6-aa
Mismatch penalty: Smaller (1)
Gap penalty: Medium (2)
Display: BLOSUM62

Alignment 18. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
    ------------------------------------------------------------

  1 MQGAQTSEDNLGSQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
              .         20          .         40          .         60

    ------------------------------------------------------------

 61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMVPPIFEGPGLEALCGETREVFG 120
              .         80          .        100          .        120

    ------------------------------------------------------------

121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
              .        140          .        160          .        180

    ------------------------------------------------------------

181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
              .        200          .        220          .        240

    ------------------------------------------------------------

241 SLQDQAKVAKVAPLKEFPASTISHPDSGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
              .        260          .        280          .        300

    ------------------------------------------------------------

301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
              .        320          .        340          .        360

    ------------------------------------------------------------

361 SFYMTHGLGTLQEGTWDPCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
              .        380          .        400          .        420

    ------------------------------------------------------------

421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
              .        440          .        460          .        480

    ------------------------------------------------------------

481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
              .        500          .        520          .        540

    ------------------------------------------------------------

541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
              .        560          .        580          .        600
```

```
     ------------------------------------------------------------

601  YSCNPFAQPACPIFTQLFYRSLLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTLC 660
               .        620          .        640          .        660

     ------------------------------------------------------------

661  TNFRSLAIDKGVLTAKEAKVVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720
               .        680          .        700          .        720

     ------------------------------------------------------------

721  KIKNRVVFTGENAALQNSFIKSTTRRENYIINGPYMKFLNTYHKTLFPDTKLSSLYLWHN 780
               .        740          .        760          .        780

     ------------------------------------------------------------

781  FSRRRSVPVPSGASAEEYSDLALFVDGGSRAHEESNVIDVVPGNLVTYAKQRLNNAILKA 840
               .        800          .        820          .        840

     ------------------------------------------------------------

841  CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900
               .        860          .        880          .        900

     ------------------------------------------------------------

901  LSQVCKARPVVTLPVTINKYTGVNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960
               .        920          .        940          .        960

     ------------------------------------------------------------

961  GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLVD 1020
               .        980          .       1000          .       1020

     ------------------------------------------------------------

1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAARDFLEGVWGGPGA 1080
               .       1040          .       1060          .       1080
                                     .         20          .
   1 -------------------STAVAQSATPSVSSSISSLRAATSGATAAA             30
                         xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
1081 AQDNFISVAEPVSTASQASAGLLLGGGGQGSGGRRKRRLATVLPGLEV*             1129
               .       1100          .       1120
```

% Identity = 0.0 (0/1129) % Homology = 0.0 (0/1129) % Total = 0.0 (0/1129)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

```
Protein  sequence      177  aa   MARRLPKPTLQG ... DTAPRGARKKQ*

Protein  sequence      221  aa   MNLAIALDSPHP ... LASLTDFLKSL*

  Method:            Diagonals (BLOSUM62)
  Layout:            Standard
  Block Length ≤:    6-aa
  Mismatch penalty:  Smaller (1)
  Gap penalty:       Medium  (2)
  Display:           BLOSUM62
```

Alignment 19. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
    1 MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
      M    +    +     L A +    P        L    + P++   R A+ +   VF+ +     E +
    1 MNLAIALDSPHPGL=ASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRTPV==EAW 57
                .          20          .         40          .
                .         80           .        100         .
   61 VQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAAASAGTGAL-ASSAPS 119
      V  +     +   +      RA  A              A  + +    A+       L A  A
   58 VAPSPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACR 117
       60          .         80          .         100         .
     120           .         140         .         160         .
  120 TAVAQSATPSVSSSISSLRAATSGATAAASAAAAVD--TGSGGGGQPHDTAPRGARKKQ* 177
      TA         V    IS LRA  +       S     V   T           D     A
  118 TADIGGKDTHVRLIISVLRAVYNDHYDYWSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGA 177
      120          .         140         .         160         .

      ---------------------------------------------

  178 IAHYLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL*                  221
      180          .         200         .         220

  % Identity =  14.7 (33/224)       % Homology =   8.0 (18/224)       % Total =  22.8 (51/224)

///
```

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:22:01 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

```
Protein  sequence      346  aa   MLSGNAGEGATA ... FCEELLNKRVA*

Protein  sequence      221  aa   MNLAIALDSPHP ... LASLTDFLKSL*
```

```
Method:            Diagonals (BLOSUM62)
Layout:            Standard
Block Length ≤:    6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:       Medium  (2)
Display:           BLOSUM62
```

Alignment 20. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
    1 MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSH 60
      M    A +     G  A+           +   +   +   +P + T   +    P  +
    1 MNLAIALDSPHP=GL=ASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRT=PVEAW 57
                .          20          .         40          .
                .         80          .        100          .        120
   61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
       A +        +  +     Y           P         L   +        +   A
   58 VAPSPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACR 117
       60          .         80          .        100          .
                .        140          .        160          .        180
  121 GLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYRKDI 180
            G      L   S LRA Y       + R R           +    D       +
  118 TADIG=GKDTHVRLII=SVLRAVYNDHYDYWS=RLRVVLCYTVVFAVRNYLDDHKSAAFV 174
      120           .         140          .        160          .
                .        200          .        220          .        240
  181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
       G       L   L   R   +                T
  175 LGAIAHYLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL*============= 221
         180          .        200          .        220
                .        260          .        280          .        300
  241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300

      ============================================================

                .        320          .        340
  301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*               346

      ================================================
```

```
% Identity =   7.8 (27/346)      % Homology =   5.2 (18/346)      % Total =  13.0 (45/346)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:23:10 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

```
Protein  sequence       24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence      221  aa   MNLAIALDSPHP ... LASLTDFLKSL*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 21. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                    .         20
     1 ---AVDTGSGGGGQPHDTA-PRGARKKQ--------------------------------- 24
          A+    S    G     T  PR
     1 MNLAIALDSPHPGLASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRTPVEAWVAP 60
               .         20          .         40          .         60

       ------------------------------------------------------------

    61 SPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACRTAD 120
               .         80          .        100          .        120

       ------------------------------------------------------------

   121 IGGKDTHVRLIISVLRAVYNDHYDYWSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGAIAH 180
               .        140          .        160          .        180

       ----------------------------------------

   181 YLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL*                     221
               .        200          .        220
```

```
  % Identity =   2.7 (6/221)       % Homology =   0.5 (1/221)       % Total =   3.2 (7/221)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:23:29 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

```
Protein  sequence       30  aa   STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence      221  aa   MNLAIALDSPHP ... LASLTDFLKSL*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 22. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                  .         20         .
     1 ST-AVAQ-SATPSVSSSISSLRAATSGATAAA---------------------------- 30
          A+A  S   P ++S      R      +
     1 MNLAIALDSPHPGLASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRTPVEAWVAP 60
               .         20         .         40         .         60

       ------------------------------------------------------------

    61 SPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACRTAD 120
               .         80         .        100         .        120

       ------------------------------------------------------------

   121 IGGKDTHVRLIISVLRAVYNDHYDYWSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGAIAH 180
               .        140         .        160         .        180

       ----------------------------------------

   181 YLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL*                    221
               .        200         .        220

  % Identity =   2.7 (6/221)       % Homology =   1.8 (4/221)       % Total =   4.5 (10/221)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 11:36:21 AM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

```
Protein  sequence      177  aa   MARRLPKPTLQG ... DTAPRGARKKQ*

Protein  sequence      222  aa   MARFIAQLLLLA ... HGVYVSGYLSQ*
```

```
Method:              Diagonals (BLOSUM62)
Layout:              Standard
Block Length ≤:      6-aa
Mismatch penalty:    Smaller (1)
Gap penalty:         Medium  (2)
Display:             BLOSUM62
```

Alignment 23. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, Ba-RF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .
    1 MARRLPKPTLQGRLEADFPDSPL-LPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEE 59
      MAR + +  L     A         L +   L         +  E + S+       +
    1 MARFIAQLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60
                .         20          .         40          .         60
     60         .         80          .        100          .
   60 YVQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAAASAGTGALASSAPS 119
            +                 A          + +              T        S
   61 GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNYLCRMKLGETEVTKQEHLS 120
                .         80          .        100          .        120
    120         .        140          .        160          .
  120 TAVAQSATP-SVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*- 177
          + +  S  S         T   T  A     V      G +P  TA  G   K+
  121 VVKPLTLSVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180
                .        140          .        160          .        180

      ------------------------------------------

  181 GSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVYVSGYLSQ*                    222
                .        200          .        220
```

```
 % Identity =  10.8 (24/222)       % Homology =   6.3 (14/222)       % Total =  17.1 (38/222)
```

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

```
Protein  sequence      346  aa   MLSGNAGEGATA ... FCEELLNKRVA*

Protein  sequence      222  aa   MARFIAQLLLLA ... HGVYVSGYLSQ*

   Method:           Diagonals (BLOSUM62)
   Layout:           Standard
   Block Length ≤:   6-aa
   Mismatch penalty: Smaller (1)
   Gap penalty:      Medium  (2)
   Display:          BLOSUM62
```

Alignment 24. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, Ba-RF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
     1 MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSH 60

       ============================================================

                .         80          .        100          .        120
    61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
                                                                  A F
     1 ============================================================MARF 4

                .        140          .        160          .        180
   121 GLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYRKDI 180
         L              +     + V    S    +     PE +      L PGE+  L  +
    .5 IAQLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLIGRMH 64
             .         20          .         40          .         60
                .        200          .        220          .        240
   181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
         +        +       R      +           G   Y       G        E     V
    65 HDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDG=NYLCRMKLGETEVTKQE=HLSV= 121
             .         80          .        100          .         120
                .        260          .        280          .        300
   241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300
          P    S   + ++Q   P   +         V      P          G   E    P       G
   122 =VKPLTLSVHSERSQ=FPDFSVLTVTCTVNAFPHPHVQWLMPEG=VE=PAP=TAANGGVM 176
                .         140          .        160               .
                .        320          .        340
   301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*                 346
       + K     + A   ++   K     P           +       +       *
   177 KEKDGSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVYVSGYLSQ*                 222
        180           .        200          .        220

  % Identity =   9.5 (33/346)       % Homology =   5.5 (19/346)       % Total =  15.0 (52/346)

///
```

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

```
Protein  sequence       24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence      222  aa   MARFIAQLLLLA ... HGVYVSGYLSQ*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 25. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, Ba-RF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 MARFIAQLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60
                ·         20          ·         40          ·         60

      ------------------------------------------------------------

   61 GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNYLCRMKLGETEVTKQEHLS 120
                ·         80          ·        100          ·        120

      ------------------------------------------------------------

  121 VVKPLTLSVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180
                ·        140          ·        160          ·        180
                                      ·         20
    1 ------------------AVDTGSGGGGQPHDTAPRGARKKQ                    24
                              G      + H      G    +
  181 GSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVYVSGYLSQ*                    222
                ·        200          ·        220

  % Identity =   1.4 (3/222)       % Homology =   0.9 (2/222)       % Total =   2.3 (5/222)

///
```

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

```
Protein  sequence       30  aa  STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence      222  aa  MARFIAQLLLLA ... HGVYVSGYLSQ*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 26. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, Ba-RF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                  .        20        .
     1 STAV-AQSATPSVSSSISSLRAATSGATAAA---------------------------- 30
           AQ    +   +       A  G
     1 MARFIAQLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60
               .        20         .        40         .        60

       ------------------------------------------------------------

    61 GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNYLCRMKLGETEVTKQEHLS 120
               .        80         .       100         .       120

       ------------------------------------------------------------

   121 VVKPLTLSVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180
               .       140         .       160         .       180

       -------------------------------------------

   181 GSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVYVSGYLSQ*                  222
               .       200         .       220

  % Identity =   1.8 (4/222)      % Homology =   0.9 (2/222)      % Total =   2.7 (6/222)

///
```

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

```
Protein  sequence       177  aa   MARRLPKPTLQG ... DTAPRGARKKQ*

Protein  sequence       163  aa   MVHVLERALLEQ ... LSLRCELGWCG*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 27. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
     1 MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
       M   L +  L+ +  A                    P DV R   R  LV L   F
     1 MVHVLERALLEQQSSACGLPGSSTETRPSHPCPEDP=DVSRL==RLLLVVLCVLFGLLCL 57
                .         20          .         40          .
                .         80          .        100          .        120
    61 VQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAASAGTGALASSAPST 120
       +        R  R +                +    + A   +          T    S
    58 LLI*EATMRPGRPLAGFYATLRRSFRRMSKRSKNKAKKERVPVEDRPP=TPMPTSQRLIR 116
        60          .         80          .        100          .
                .        140          .        160          .
   121 AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*    177
         A         +     R           S     D  S
   117 RNALGGGVRPDAEDCIQRFHPLEPALGVSTKNF=DLLSLRCELGWCG*=========    163
        120          .        140          .        160     .
```

```
  % Identity =  13.0 (23/177)      % Homology =   5.1 (9/177)      % Total =  18.1 (32/177)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 1:00:26 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

```
Protein  sequence      346  aa   MLSGNAGEGATA ... FCEELLNKRVA*

Protein  sequence      163  aa   MVHVLERALLEQ ... LSLRCELGWCG*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 28. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .     .  20          .         40          .
    1 MLSGNAGEGATACGGSAAAGQDLISVP-RNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFS 59
      M+      E A      S+A G     S    R +        ++           L      L
    1 MVHVL==ERALLEQQSSACGLPGSSTETRPSHPCPEDPDVSRLRLLLVVLCVLFGLLCLL 58
                    .          20          .         40          .
     60             .         80           .        100          .
   60 HQAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAY 119
         AT      P AG  A      F           A  +           T  P
   59 LI*EATMRPGRPLAGFYATLRRSFRRMSKRSKNKAKKERVPVEDRPP=TPMPTSQRLIRR 117
      60           .         80           .        100          .
    120            .        140           .        160          .
  120 FGLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYRKD 179
        L G    P      +      L       +     +       E     G
  118 NALGGGVRPDAEDCIQRFHPLEPALGVSTKNFDLLSLRCELGWCG*============== 163
      120            .        140           .        160
    180              .        200           .        220          .
  180 IAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYV 239

      ============================================================

    240              .        260           .        280          .
  240 PHPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGH 299

      ============================================================

    300              .        320           .        340
  300 HRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*               346

      ================================================
```

```
 % Identity =   7.5 (26/347)       % Homology =   2.6 (9/347)       % Total =  10.1 (35/347)
```

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

```
Protein  sequence       24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence      163  aa   MVHVLERALLEQ ... LSLRCELGWCG*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 29. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20
    1 AVDTGSGGGGQPHDTAPRGARKKQ----------------------------------- 24
       V         +    +A
    1 MVHVLERALLEQQSSACGLPGSSTETRPSHPCPEDPDVSRLRLLLVVLCVLFGLLCLLLI 60
                .         20          .         40          .         60

      ------------------------------------------------------------

   61 *EATMRPGRPLAGFYATLRRSFRRMSKRSKNKAKKERVPVEDRPPTPMPTSQRLIRRNAL 120
                .         80          .        100          .        120

      ---------------------------------------------

  121 GGGVRPDAEDCIQRFHPLEPALGVSTKNFDLLSLRCELGWCG*                   163
                .        140          .        160
```

```
 % Identity =   1.2 (2/163)      % Homology =   1.2 (2/163)      % Total =   2.5 (4/163)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 1:01:19 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

```
Protein  sequence        30  aa   STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence       163  aa   MVHVLERALLEQ ... LSLRCELGWCG*
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 30. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20         .
     1 STAVAQSATPSVSSSISSLRAATSGATAAA------------------------------ 30
         V + A      SS    L  +++    +
     1 MVHVLERALLEQQSSACGLPGSSTETRPSHPCPEDPDVSRLRLLLVVLCVLFGLLCLLLI 60
                .         20         .         40         .         60

       ------------------------------------------------------------

    61 *EATMRPGRPLAGFYATLRRSFRRMSKRSKNKAKKERVPVEDRPPTPMPTSQRLIRRNAL 120
                .         80         .        100         .        120

       -------------------------------------------

   121 GGGVRPDAEDCIQRFHPLEPALGVSTKNFDLLSLRCELGWCG*                   163
                .        140         .        160

  % Identity =   3.1 (5/163)       % Homology =   3.1 (5/163)       % Total =   6.1 (10/163)

///
```

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

```
Protein  sequence      177  aa   MARRLPKPTLQG ... DTAPRGARKKQ*

Protein  sequence      269  aa   VLGIWIYLLEML ... PHGPVQLSYYD*

   Method:            Diagonals (BLOSUM62)
   Layout:            Standard
   Block Length ≤:    6-aa
   Mismatch penalty:  Smaller (1)
   Gap penalty:       Medium  (2)
   Display:           BLOSUM62
```

Alignment 31. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 VLGIWIYLLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQNWWTLLVDLLWLLLFL 60
                ·         20          ·         40          ·         60
                                                ·         20
    1 ---------------------------------MARRLPKPTLQGRLEADFPDSPLLPKFQ 28
                                          + +GR        +    P
   61 AILIWMYYHGQRHSDEHHHDDSLPHPQQATDDSGHESDSNSNEGRHHLLVSGAGDGPPLC 120
                ·         80          ·        100          ·        120
         ·         40          ·         60          ·         80
   29 ELNQNNLPNDVFREAQRSYLVFLTSQFCYEEYVQRTFGVPRRQRAIDKRQRASVAGAGAH 88
        N              Q               +          P  Q    D
  121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQ=DPDNTDDNGPQDPDNT 179
                ·       140          ·       160          ·
         ·       100          ·       120          ·       140
   89 AHLGGSSATPVQQAQAAASAGTGALASSAPSTAVAQSATPSVSSSISSLRAATSGATAAA 148
          G      P    + +A + G G     +            P + +       +   SG
  180 DDNGPHDPLPHSPSDSAGNDG=GPPQLTEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGD 238
    180           ·       200          ·       220          ·
         ·         160          ·
  149 SAAAAVDTG-SGGGGQPHD-TAPRGARKKQ*                                177
            +   G SG GG    D    P        *
  239 PHLPTLLLGSSGSGGDDDDPHGPVQLSYYD*                                269
      240           ·       260

   % Identity =   8.9 (24/271)       % Homology =   4.4 (12/271)       % Total =  13.3 (36/271)

///
```

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 12:57:10 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

```
Protein  sequence      346  aa   MLSGNAGEGATA ... FCEELLNKRVA*

Protein  sequence      269  aa   VLGIWIYLLEML ... PHGPVQLSYYD*
```

```
Method:              Diagonals (BLOSUM62)
Layout:              Standard
Block Length ≤:      6-aa
Mismatch penalty:    Smaller (1)
Gap penalty:         Medium  (2)
Display:             BLOSUM62
```

Alignment 32. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
    1 MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSH 60

      ============================================================

                .         80          .        100          .        120
   61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
                        V     Y       +  G        AF
    1 ===============VLGIWIY=LLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQ 44
                                    .        20          .         40
                .         140         .        160          .        180
  121 GLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYRKDI 180
       ,  L       LL    + L    Y       S++    D +
   45 NWWTL=LVDLLWLLLFLAILIWMYYHGQRHSDEHHHD=DSLPHPQQATDDSGHESDSNSN 102
              .         60          .         80          .        100
                .         200         .        220          .        240
  181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
       G     +          L  + L      G     PQ        +GP              +P
  103 EGRHHLLVSGAGDGPPLCSQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLP 162
              .         120         .        140          .        160
                .         260         .        280          .
  241 -HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGH 299
        P      +              G    P  H  ++S    + G     +    E+     Q
  163 QDPDNTDDNGPQDPDNTDDNGPHDPLPHSPSDSAGNDGGPPQLTEEVENKGGDQGPPLMT 222
              .         180         .        200          .        220
    300         .         320         .        340
  300 HRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*               346
        G       S  G       PT      S S          +       *
  223 DGGGGHSHDSGHGGGDPHLPTLLLGSSGSGGDDDDPHGPVQLSYYD*               269
              .         240         .        260
```

```
 % Identity =  10.7 (37/347)       % Homology =   4.6 (16/347)       % Total =  15.3 (53/347)
```

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

```
Protein  sequence        24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence       269  aa   VLGIWIYLLEML ... PHGPVQLSYYD*
```

```
Method:            Diagonals (BLOSUM62)
Layout:            Standard
Block Length ≤:    6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:       Medium  (2)
Display:           BLOSUM62
```

Alignment 33. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 VLGIWIYLLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQNWWTLLVDLLWLLLFL 60
                ·         20          ·         40          ·         60

      ------------------------------------------------------------

   61 AILIWMYYHGQRHSDEHHHDDSLPHPQQATDDSGHESDSNSNEGRHHLLVSGAGDGPPLC 120
                ·         80          ·        100          ·        120

      ------------------------------------------------------------

  121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQDPDNTDDNGPQDPDNTD 180
                ·        140          ·        160          ·        180

      ------------------------------------------------------------

  181 DNGPHDPLPHSPSDSAGNDGGPPQLTEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGDPH 240
                ·        200          ·        220          ·        240
                       ·         20
    1 -----AVDTGSGGGGQPHDTAPRGARKKQ                                  24
            +GSGG          + +
  241 LPTLLLGSSGSGGDDDDPHGPVQLSYYD*                                   269
                ·        260
```

```
 % Identity =   1.5 (4/269)       % Homology =   1.1 (3/269)       % Total =   2.6 (7/269)
```

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

```
Protein  sequence       30  aa   STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence      269  aa   VLGIWIYLLEML ... PHGPVQLSYYD*
```

```
Method:            Diagonals (BLOSUM62)
Layout:            Standard
Block Length ≤:    6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:       Medium  (2)
Display:           BLOSUM62
```

Alignment 34. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
      ------------------------------------------------------------

    1 VLGIWIYLLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQNWWTLLVDLLWLLLFL 60
                ·         20          ·         40          ·         60

      ------------------------------------------------------------

   61 AILIWMYYHGQRHSDEHHHDDSLPHPQQATDDSGHESDSNSNEGRHHLLVSGAGDGPPLC 120
                ·         80          ·        100          ·        120

      ------------------------------------------------------------

  121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQDPDNTDDNGPQDPDNTD 180
                ·        140          ·        160          ·        180

    1 ----------------------------------------------------------ST 2

  181 DNGPHDPLPHSPSDSAGNDGGPPQLTEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGDPH 240
                ·        200          ·        220          ·        240
              .          20          .
    3 AVAQSATPSVS-SSISSLRAATSGATAAA                                  30
             S S             +
  241 LPTLLLGSSGSGGDDDDPHGPVQLSYYD*                                  269
                ·        260
```

```
% Identity =   0.7 (2/269)      % Homology =   0.4 (1/269)      % Total =   1.1 (3/269)
```

///

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

```
Protein  sequence       177  aa   MARRLPKPTLQG ... DTAPRGARKKQ*

Protein  sequence       144  aa   MEHDLERGPPGP ... LGIVLFIFGCLL
```

```
Method:            Diagonals (BLOSUM62)
Layout:            Standard
Block Length ≤:    6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:       Medium  (2)
Display:           BLOSUM62
```

Alignment 35. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
    1 MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
      M   L +     R     P          L    L    +     S          Y
    1 MEHDLERGPPGPRRPPRGPPLSSSLGLALLLLL=LALLFWLYIVMSDWTGGALLVLYSFA 59
                .         20          .         40          .
                .         80          .        100          .        120
   61 VQRTFGVPRRQRAIDKRQRASVAGAGAHAHLGGSSATPVQQAQAAASAGTGALASSAPST 120
      +     +       I +R      GA     L     T          S+ + AL+  +P T
   60 LMLIIIILIIF==IFRRDLLCPLGALCILLLMSKYYTLCPTPPFPYSSFSNALSPLSPVT 117
     60          .           80          .        100          .
                 .       140          .        160          .
  121 AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*     177
       +   A  ++       L
  118 LLLI=ALWNLHGQALFLGIVLFIFGCLL============================     144
      120           .       140
```

```
  % Identity =  11.9 (21/177)      % Homology =   6.8 (12/177)      % Total =  18.6 (33/177)
```

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

```
Protein  sequence      346  aa   MLSGNAGEGATA ... FCEELLNKRVA*

Protein  sequence      144  aa   MEHDLERGPPGP ... LGIVLFIFGCLL
```

```
Method:            Diagonals (BLOSUM62)
Layout:            Standard
Block Length ≤:    6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:       Medium  (2)
Display:           BLOSUM62
```

Alignment 36. Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20          .         40          .         60
    1 MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSH 60
      M       G           G  L S      + LL   L       +     A L  +S
    1 MEHDLER=GPPGPRRPPR=GPPLSSSLGLALLLLLLALLFWLYIVMSDWTGGALLVLYSF 58
                .          20          .         40          .
                .         80          .        100          .        120
   61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
        +                                                +   +       P
   59 ALMLIIIILIIFIFRRDLLCPLGALCILLLMSKYYTLCPTPPFPYSSFSNALSPLSPVTL 118
      60          .         80          .        100          .
                .        140          .        160          .        180
  121 GLPGLFGPPPPCLLTTDSHLRADYVPAPSRSNKRKRDPEEDEEGGGLFPGEDATLYRKDI 180
       L  L+      L            +
  119 LLIALWNLHGQALFLGIVLFIFGCLL================================== 144
     120          .        140
                .        200          .        220          .        240
  181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240

      ============================================================

                .        260          .        280          .        300
  241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300

      ============================================================

                .        320          .        340
  301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA*               346

      ============================================
```

```
  % Identity =   4.3 (15/346)      % Homology =   2.3 (8/346)      % Total =   6.6 (23/346)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 12:55:12 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

```
Protein  sequence        24  aa   AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein  sequence       144  aa   MEHDLERGPPGP ... LGIVLFIFGCLL
```

```
Method:             Diagonals (BLOSUM62)
Layout:             Standard
Block Length ≤:     6-aa
Mismatch penalty:   Smaller (1)
Gap penalty:        Medium  (2)
Display:            BLOSUM62
```

Alignment 37. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                 .         20
     1 AVDTGSGGGGQPHDTAPRGARKKQ----------------------------------- 24
               G    P     PRG
     1 MEHDLERGPPGPRRP=PRGPPLSSSLGLALLLLLLALLFWLYIVMSDWTGGALLVLYSFA 59
                 .          20          .         40          .

       ------------------------------------------------------------

    60 LMLIIIILIIFIFRRDLLCPLGALCILLLMSKYYTLCPTPPFPYSSFSNALSPLSPVTLL 119
      60          .         80          .        100          .

       -------------------------

   120 LIALWNLHGQALFLGIVLFIFGCLL                                    144
     120          .        140

  % Identity =   3.4 (5/145)       % Homology =   0.0 (0/145)       % Total =   3.4 (5/145)
```

///

DNA Strider 1.4f6 ### Wednesday, July 18, 2007 12:55:30 PM (US Letter @ 100%)

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

```
Protein  sequence       30  aa   STAVAQSATPSV ... LRAATSGATAAA

Protein  sequence      144  aa   MEHDLERGPPGP ... LGIVLFIFGCLL
```

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty: Smaller (1)
Gap penalty:      Medium  (2)
Display:          BLOSUM62
```

Alignment 38. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```
                .         20         .
     1 STAVAQSATPSVSSSISSLRAATSGATAAA------------------------------ 30
             +   P            ++S   A
     1 MEHDLERGPPGPRRPPRGPPLSSSLGLALLLLLLALLFWLYIVMSDWTGGALLVLYSFAL 60
                .         20         .         40         .         60

       ------------------------------------------------------------

    61 MLIIIILIIFIFRRDLLCPLGALCILLLMSKYYTLCPTPPFPYSSFSNALSPLSPVTLLL 120
                .         80         .        100         .        120

       ------------------------

   121 IALWNLHGQALFLGIVLFIFGCLL                                     144
                .        140

  % Identity =   2.1 (3/144)       % Homology =   2.1 (3/144)       % Total =   4.2 (6/144)
```

///